CS 468 NOTES: DIFFERENTIAL GEOMETRY FOR COMPUTER SCIENCE

ARUN DEBRAY
MAY 20, 2013

Contents

1. Overview: 4/1/13

"The first lecture of any CS course is always kind of propaganda."

Classical Euclidean geometry is well-studied, but is very bad for surfaces that aren't circles, points, or lines. Thus, differential geometry was developed to study smooth surfaces.

Defining a smooth surface requires a bit of thought, and can't really be done with Euclidean geometry.

Definition. A function $f : \mathbb{R} \to \mathbb{R}$ is C^∞ if the k^{th} derivative of f exists and is continuous for all $k \in \mathbb{N}$.

However, there's a bit of divergence for smooth surfaces: it's simply not that simple. Consider two functions $f_1, f_2 : \mathbb{R} \to \mathbb{R}^2$ given by $f_1(t) = (t, 2t)$ and

$$f_2(t) = \begin{cases} (t, 2t), & t < 1 \\ \left(2\left(t - \frac{1}{2}\right), 4\left(t - \frac{1}{2}\right)\right), & t \geq 1. \end{cases}$$

f_1 and f_2 trace out the same curve $y = 2x$, but the velocity isn't continuous in f_2. Thus, f_2 isn't C^∞, but f_2 isn't, so is this a smooth curve? Thus, one might propose that a smooth surface is one that admits a smooth parameterization.

But that's still not quite enough: consider $f(t) = (t^2, t^3)$, which has a cusp at $t = 0$. Thus, it doesn't seem sensible to call it smooth, but f is C^∞!

Moreover, a smooth surface might not even be the graph of a functon, such as the double torus: it's a set of points in $\mathbb{R}^3$ (also called ambient space). However, its interior isn't a surface. How should that be distinguished?

It's time to bring in another idea from high school calculus: a smooth function is locally linear, even though it may be more interesting globally. Thus, smooth surfaces could be defined to be those which locally look like planes. The idea of coordinate charts comes in here. This will all be formalized in a future lecture.

Two basic principles of smooth surfaces are Gaussian and mean curvature, which are both completely independent of parameterization. One can also consider flows and vector fields, which allow the finding of interesting feature points (e.g. extrema) of curves, which leads to Morse Theory. Ideally, the goal is to find points that are invariant under certain transformations.

One also consideres distances in differential geometry. Dijkstra's algorithm can be used to find an approximation of distance on a triangulated version of a curve.

These notions can be generalized to manifolds (k-dimensional objects sitting in n-dimensional space). This course will consider mostly surfacs in 3-space ($k = 2$, $n = 3$), since they're often easier to deal with than volumes.

Another viewpoint is that of Riemannian geometry, which thinks only about angles and distances, as if from the point of view of an ant on the surface. Thus, depending on the curvature of the surface, lines that are locally parallel might converge or diverge in the cases of positive and negative curvature, respectively. This can be applied in the study of map projections: are two lines on a map the same length on the sphere?

Another subject of study is that of not-quite-smooth surfaces. Some might have boundaries or sharp edges. This doesn't really matter in the discrete case.

Geometry can also show up in surprising places: if one considers the double pendulum, the state space is given by the two angles θ_1 and θ_2, and since each of these is given by a circle, the total space is their product, a torus. This is useful for simulation, as it simplifies the physics.

Turning to the discrete case, the typical geometric input is a big set of points produced by (say) a laser scanner or something such as a self-driving car. The set of points is processed into a triangle mesh.[1] A mesh is formally a collection of vertices V and triangles T with some extra conditions.

A simple counterexample to show that the conditions are necessary is the nonmanifold edge. Notice that three faces touch at a vertex. Thus, a manifold requires that each edge is incident to at most 2 faces, and the two faces incident to a vertex form a closed or open fan. These two conditions are sufficient for surface-type analysis.

One concern is that of invalid or bad meshes. Topologically it might be fine, but the parameterization matters: most techniques work betteron meshes with uniform lines and angles. Another concern is how functions are translated to the mesh. One value per vertex seems reasonable until it loses important information and also ignores the continuous properties. Choosing approximations of functions can get quite complicated.[2] Under stricter assumptions the values might converge.

The way to understand this is to consider discrete versus discretized geometry. The former starts with discreteness and works towards the continuous geometry, and the latter develops discretized analogues of continuous concepts. More recently, there has been discrete differential geometry, which is a discrete parallel to differential geometry, defining (for example) curvature in a way that leads to similar or parallel theorems. This is much nicer than in the discretized case, where theorems generally don't transfer. Structure preservation, keeping properties of the continuous surface true in the discrete one, is an important goal.

There are many applications of computational differential geometry:

- Computer graphics: transferring and exploiting patterns, editing models
- Computer vision (CS 231): navigation, facial recognition, reconstructing or segmenting surfaces (for 3D printing)
- Medical imaging
- Manufacturing: scanning, detection of defects
- Architecture, especially considering more recent trends in curved surfaces.

Here are some example tasks:

- Deformation of some object (which has use in animation)
- Expressive rendering (art!)
- Shape collection analysis: finding corresponding points in several related models. This is difficult when they aren't related by rigid motions.
- Relatedly, finding mappings between surfaces, perhaps with some additional constrait such as preserving a function on a surface. This is mostly an unsolved problem.
- This leads to deformation transfers, such as mapping an actor's gestures onto a computer-generated model.
- Simulation is a big one:
  - Viscous threads: some liquids, such as honey, behave like curves in certain cases.
  - Cloth simulation, where some cloths bend much more easily than they stretch.
- Scientific visualization, related to many of the above.

All of these diverse applications have the theory in common. There is much more to this stuff than one class can really cover, though.

[1] This step could easily be a course's worth of material.

[2] See *Convergence Analysis of Discrete Differential Geometry Operators over Surfaces*, by Zhiqiang Xu, Guoliang Xu, and Jia-Guang Sun. This paper concludes there is no discrete scheme such that Gaussian curvature, mean curvature, and the Laplace-Beltrami operator converge over a triangulated surface, and that there always must be some weird counterexample for any scheme.

2. Curves: 4/3/13

"It's my turn to give the lecture today. It's sort of a good-cop bad-cop thing, but it's up to you to figure out which is which." – Adrian Butscher

Curves are a good place to start with: they're the one-dimensional version of surfaces, and are thus somewhat easier to understand.

Definition. A parameterized differentiable curve is a differentiable map $\gamma : I \to \mathbb{R}^n$, where $I = (a, b) \subset \mathbb{R}$. I is called the parameter domain, and the curve itself can be called the image or trace of γ. Additionally, γ can be written $\gamma(t) = (\gamma_1(t), \ldots, \gamma_n(t))$, and these latter functions are called the component functions.

We will be most concerned with the case $n = 3$, of course. Note that the image of γ is the geometric object, but the function is how it's accessed. $\gamma(t)$ should be thought of as a point on the curve at time t, and as time progresses, the curve moves. Thus, there are multiple ways to parameterize a curve (so multiple such γ exist), since one might traverse a curve with different speeds.

Definition. The velocity of a curve is the vector $V(t) = \frac{\mathrm{d}\gamma}{\mathrm{d}t}$. Thus, if $V(t) = (V_1(t), V_2(t), V_3(t))$, then $V_i(t) = \frac{\mathrm{d}\gamma_i}{\mathrm{d}t}$.

This is a vector originating at $\gamma(t)$ and is tangent to the curve, since it is the limit of a sequence of secant lines $\gamma(t + h)$ as $h \to 0$. From this the formal definition can be recovered: $V(t) = \lim_{h\to 0}(\gamma(t + h) - \gamma(t))/h$. This wasn't very rigorous, but it can be made to be so.

Definition. The acceleration is the rate of change of the velocity vector: $A(t) = \frac{\mathrm{d}^2\gamma}{\mathrm{d}t^2}$.

This doesn't have such a nice geometric visualization such as the tangency of the velocity.

In the case of constant-velocity curves (i.e. each component of γ has a constant derivative), then γ is linear in t:

$$\gamma(t) = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} t + \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}.$$

Thus, these curves are easy to understand: they are just parameterizations of lines. However, it gets slightly more interesting in the case of constant speed. Here $\|V(t)\| = \sqrt{\gamma_1'(t)^2 + \gamma_2'(t)^2 + \gamma_3'(t)^2}$ is made to be constant. Thus, the square of the speed is constant:

$$\begin{aligned} \left\| \frac{\mathrm{d}\gamma}{\mathrm{d}t} \right\|^2 &= \dot\gamma_1^2 + \dot\gamma_2^2 + \dot\gamma_3^2 \\ \implies 0 &= 2\dot\gamma_1\ddot\gamma_1 + 2\dot\gamma_2\ddot\gamma_2 + 2\dot\gamma_3\ddot\gamma_3. \end{aligned}$$

Now, we can pull an inner product out of this:

$$\begin{aligned} &= 2\left\langle \begin{pmatrix} \dot\gamma_1 \\ \dot\gamma_2 \\ \dot\gamma_3 \end{pmatrix}, \begin{pmatrix} \ddot\gamma_1 \\ \ddot\gamma_2 \\ \ddot\gamma_3 \end{pmatrix} \right\rangle \\ &= 2\left\langle V(t), A(t) \right\rangle. \end{aligned}$$

Thus, if a curve has constant speed, then its velocity and acceleration are perpendicular.

Example 2.1. Many of these examples might be familiar:

- If $\gamma(t) = (\cos t, \sin t)$, then γ traces out a circle with constant speed (since $\cos^2 t + \sin^2 t = 1$).
- The helix is given by $\gamma(t) = (\cos t, \sin t, t)$.
- Curves can have self-intersection, as when $\gamma(t_1) = \gamma(t_2)$ for $t_1 \neq t_2$. This is absolutely OK.
- Curves can have kinks or cusps, such as $\gamma(t) = (t^3, t^2)$. Correspondingly, a smooth curve can have a non-smooth parameterization; be careful!

It is important to discuss changes in parameterization. Suppose $\gamma : I \to \mathbb{R}^3$ and $\varphi : I \to I$ is a smooth bijection (i.e. there exists a smooth inverse, or the function is injective and surjective). Since φ is a bijection on an interval, then it is monotonic, and φ' is thus nonzero.

Using φ, it is possible to give a reparameterization of the curve: $\widetilde{\gamma}(s) = \gamma \circ \varphi(s)$. This curve has the same trace, but its velocity can change. Using the Chain Rule, which is a very important technique in differential geometry,

$$\frac{\mathrm{d}\widetilde{\gamma}}{\mathrm{d}s} = \frac{\mathrm{d}(\gamma \circ \varphi(s))}{\mathrm{d}t}\frac{\mathrm{d}\varphi}{\mathrm{d}s} = V(\varphi(s))\frac{\mathrm{d}\varphi}{\mathrm{d}s} = V(t)\frac{\mathrm{d}\varphi}{\mathrm{d}s}.$$

Since φ isn't guaranteed to be the identity, the velocity can be different.

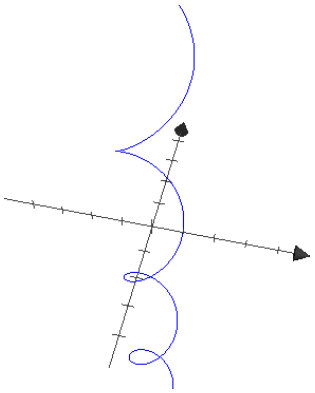

FIGURE 1. A helix, given by $\gamma(t) = (\cos t, \sin t, t)$.

It is useful to be able to compute the length of a curve, which is a nice application of integral calculus. Suppose $\gamma : [a, b] \to \mathbb{R}^3$ is a smooth curve and partition $I = [a, b]$ as $I = [t_0, t_1] \cup \cdots \cup [t_{n-1}, t_n]$, where $t_0 = a$ and $t_n = b$. Then, the arc length can be approximated as the sum of the lengths of the lines connecting $\gamma(t_i)$ and $\gamma(t_{i+1})$, so γ is approximated by a piecewise linear curve. Thus, if ℓ is the arc length of γ, then

$$\ell \approx \sum_{i=1}^{n} \|\gamma(t_i) - \gamma(t_{i-1})\| = \sum_{i=1}^{n} \|\dot\gamma(t_i^*)\Delta t_i\| = \sum_{i=1}^{n} \|\dot\gamma(t_i^*)\|\Delta t_i.$$

t_i^* is given by the Mean Value Theorem.[3] Then, taking the limit gives the exact value, since the above is a Riemann sum:

$$\ell = \lim_{n\to\infty} \sum_{i=1}^{n} \|\dot\gamma(t_i^*)\|\Delta t_i = \int_a^b \|\dot\gamma(t)\| \, \mathrm{d}t.$$

The arc length is a geometrical quantity, and it is invariant of parameterization. This is one of the key points of the lecture, and is relatively surprising, since the specific paramaterization of γ was used to compute the length. Here's why: suppose $\phi : [a, b] \to [a, b]$ is an (orientation-preserving) diffeomorphism such that $\phi(a) = a$ and $\phi(b) = b$. Then, let $\widetilde\gamma(s) = \gamma(\phi(s))$, and let ℓ be the length of γ and $\widetilde\ell$ be the length of $\widetilde\gamma$. Then, using the Chain Rule again,

$$\begin{aligned}
\widetilde\ell &= \int_a^b \left\| \frac{\mathrm{d}(\gamma \circ \phi(s))}{\mathrm{d}s} \right\| \mathrm{d}s \\
&= \int_a^b |\phi'(s)| \left\| \frac{\mathrm{d}\gamma}{\mathrm{d}t} \circ \phi(s) \right\| \mathrm{d}s \\
&= \int_a^b |\phi' \circ \phi^{-1}(t)| \left\| \frac{\mathrm{d}\gamma(t)}{\mathrm{d}t} \right\| \frac{\mathrm{d}t}{|\phi' \circ \phi^{-1}(t)|}.
\end{aligned}$$

The above might seem scary, but it's just a substitution using the formula for the derivative of ϕ^{-1}.

$$= \int_a^b \left\| \frac{\mathrm{d}\gamma(t)}{\mathrm{d}t} \right\| \mathrm{d}t = \ell.$$

This is illustrative of a more general trend in differential goemetry, in which various geometric quantities are calculated via some parameterization. However, it is then necessary to show that such a calculation is independent of parameterization. However, in most cases, it will be much more complicated, and the full rigorous proof won't always be given.

Example 2.2. Suppose γ is a unit speed parameterization of a curve, so that $\|\dot\gamma\| = 1$. This is nice, because there is a closed-form expression, which isn't always the case: $\ell = \int_0^T \|\dot\gamma\| \, \mathrm{d}t = \int_0^T \mathrm{d}t = T$. Intuitively, this makes sense given that the curve must be a line.

More generally, there can exist parameterizations of nonlinear γ such that the length on $[0, T]$ is just T. This is very nice in terms of calculating arc length, and also implies that the velocity and acceleration are orthogonal.

[3] If you don't remember what that is, it states that if f is a differentiable curve on an interval $[a, b]$, then there exists a $z \in [a, b]$ (called "zed" by the instructor despite his American accent) such that $(f(y) - f(x))/(y - x) = f'(z)$: the slope of the tangent line is equal to the slope of the secant line.

It happens that every curve admits such a parameterization; see the handout for the proof. The existence of this parameterization is part of what makes curves so special.

Curves also have a property called curvature, which puts the acceleration to good use, albeit after modifying it a bit.

Definition. The geodesic curvature vector is the normal component of acceleration vector, normalized by the velocity vector (so that it's parameterization-independent): $\mathbf{k}_\gamma = [\ddot{\gamma}]^\perp / \|\dot{\gamma}\|^2$. Then, the geodesic curvature is $k_\gamma = \|\mathbf{k}_\gamma\|$.

In the arc-length parameterization, one has $\mathbf{k}_\gamma = [\ddot{\gamma}]^\perp$. However, in the general case, this can be computed by subtracting off the tangential part:

$$\mathbf{k}_c = \frac{1}{\|\dot{\gamma}\|^2}\left(\ddot{\gamma} - \frac{\langle \dot{\gamma}, \ddot{\gamma}\rangle \dot{\gamma}}{\|\dot{\gamma}\|^2}\right) = \frac{1}{\|\dot{\gamma}\|^2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\dot{\gamma}}{\|\dot{\gamma}\|}\right).$$

Exercise 2.1. Why does the second equality hold?

Notice that the latter quantity contains a unit vector, called the unit tangent vector, T.

Another helpful concept is the Frenet frame. Assume without loss of generality γ is parameterized by arc length.

Definition. A frame on a curve γ is a set of 3 orthogonal coordinates at every point of the curve.

The Frenet frame is given as follows:

- The first coordinate vector is the unit tangent vector: $T(s) = \dot{\gamma}(s)$. This doesn't need to be normalized, since we are in the arc-length parameterization.
- At a point with nonzero curvature, the second choice is the unit normal vector $N(s) = \dot{T}(s)/\|\dot{T}(s)\|$.
- The last choice is the binormal vector $B(s) = T(s) \times N(s)$.

Notice this isn't defined where the curvature is zero (e.g. a straight line). The osculating plane is the span of $T(s)$ and $N(s)$.

There is a system of equations called the Frenet formulas that link the curvature to the changes in the vectors:

$$\begin{aligned}\dot{T}(s) &= k_\gamma(s)N(s)\\ \dot{N}(s) &= -k_\gamma(s)T(s) - \tau_\gamma(s)B(s)\\ \dot{B}(s) &= \tau_\gamma(s)B(s).\end{aligned}$$

Here, $\tau_\gamma(s)$ is the torsion of the curve: $\tau_\gamma(s) = -\left\langle \dot{N}(s), B(s)\right\rangle$.

So why are these important? In differential geometry, there are local theorems, which hold true in some coordinate neightborhood, and global theorems, which hold true for the entire curve. The Frenet frame allows one to prove both a local and a global theorem.

Locally, k and $\dot{k}$ determine the amount that the curve turns in the osculating plane, and τ and k determine how much the curve lifts out of the osculating plane into the normal direction. Thus, the local behavior of the curve is characterized in terms of the curvature and torsion.

Globally, we have:

Theorem 2.1 (Fundamental Theorem of Curves)**.** *Suppose $k : I \to \mathbb{R}^+$ and $\tau : I \to \mathbb{R}$ are differentiable functions. Then, there exists a curve $\gamma : I \to \mathbb{R}^3$ such that s is the arc-length, $k(s)$ is the geodesic curvative, and $\tau(s)$ is the torsion, and any other such curve differs from γ by a rigid motion.*

This theorem states that the curvature and torsion functions uniquely determine a curve, up to location in $\mathbb{R}^3$ and rotations.

3. Discrete Curves: 4/8/13

"I wish I were the derivative of your parameterization so I could lie tangent to your curves."

Reviewing the material from last week, there was an extremely useful theorem that was used implicitly several times.

Exercise 3.1. Take $v(t) : \mathbb{R} \to \mathbb{R}^n$ such that $\|v(t)\| = 1$ for all t. Show that $\langle v(t), v'(t)\rangle = 0$ for all t.

Solution. $1 - \|v\|^2 = \langle v, v\rangle$, and take the derivative: $0 = \frac{\partial}{\partial t}\langle v, v\rangle = 2\langle v, v'\rangle$. ⊠

This is used to construct the Frenet frame and lots of useful proofs. The Frenet fram is a set of three orthogonal vectors: the tangent vector $\mathbf{T}$, which is an approximation of the curve, the normal vector $\mathbf{N}$, which along with the tangent approximates the plane the curve lies in, and the binormal $\mathbf{B}$. These lead to the ideas of curvature and torsion. Curvature seems relatively intuitive, but torsion is a bit odder: a planar curve has no torsion, and in general the torsion approximates how much the curve leaves the plane given by $\mathbf{T}$ and $\mathbf{N}$ at any given point.

On to the discretization. The most obvious approximation is to give a curve by a piecewise-smooth approximation, such as in a Bézier curve. This is already sufficient for many applications in compuer graphics, but for differential geometry it has issues: calculus doesn't work at these discontinuities.

A cubic Bézier curve has arc length $s = \int_{t_0}^{t_1} \sqrt{\gamma_x'^2 + \gamma_y'^2 + \gamma_z'^2}\,\mathrm{d}t$, which is actually very hard and sometimes impossible to obtain a closed-form solution for. Thus, getting things such as the arc-length parameterization aren't quite viable. This reflects a more general truth that closed-form solutions of things in general are quite difficult, messy, or impossible.

Thus, one might want to try something different than Bézier curves. An equally good approximation is a picewise linear one: it can approximate a curve just as well as a Bézier curve, though it needs more points to do so. That's mostly a matter of computational power. However, a piecewise-linear curve has bad curvature: on the lines, the curvature is zero, and at the vertices, the curvature is infinite.

Trying another approach, one could try the finite difference trick, which results in theorem of the form $h \to 0$ implies something about an approximation, such as $f'(x) \approx (f(x+h) - f(x))/h$. But there are many cases where things don't converge well, and in CS one might not be able to send $h \to 0$ all that cleanly. Thus, the goal is to learn something about the continuous structure while still in the discrete world. There are several good models, but no single model works well for all applications.

Working with planar curves, there is a map called the Gauss map from the curve to S^1, in which the unit normal is sent to the point on the circle corresponding to its direction. On the plane, there is also an angular formula: given canonical xy-axes in $\mathbb{R}^2$, then let $\theta(s)$ be the angle beteween the unit tangent and the x-axis. Then, $T(s) = (\cos(\theta(s)), \sin(\theta(s)))$, and take the derivative: $T'(s) = \theta'(s)(-\sin(\theta(s)), \cos(\theta(s)))$. This can be thought of as a signed curvature $\kappa(s) = \theta'(s)$ and a unit normal $N(s) = (-\sin(\theta(s)), \cos(\theta(s)))$. Then, this can be used to obtain things such as the turning number of the Gauss map (how many times the curve loops around, informally). This depends in θ, but this can be recovered by integrating the curvature.

Theorem 3.1 (Turning Number)**.**

$$\int \kappa(s)\,\mathrm{d}s = 2\pi k.$$

This is a global theorem, about the whole curve. But it has applications in discrete geometry: a piecewise linear approximation still has a Gauss map, and this can be used to recover the turning number. If Γ is a neighborhood of the curve as two segments meeting at an angle θ with lengths ℓ_1 and ℓ_2, then $\kappa \approx \theta/(\ell_1 + \ell_2)$, representing the total change in curvature. The angle is given by $\theta = \int_\Gamma \kappa(s)\,\mathrm{d}s$ more generally.

However, if one adds more points to the same curve, the curvature increases, since the segments among which the curvature is calculated are shorter. However, the integrated curvature is the same, as is θ.[4]

Theorem 3.2 (Discrete Turning Number)**.** *If Γ is a piecewise linear curve, where the piecewise components are $\Gamma_1, \ldots, \Gamma_n$, then*

$$\int_\Gamma \kappa\,\mathrm{d}s = \sum_{i=1}^{n} \int_{\Gamma_i} \kappa(s)\,\mathrm{d}s = \sum_{i=1}^{n} \theta_i = 2\pi k.$$

This is a kind of silly theorem, but it's remarkable in that it parallels the continuous definition exactly at all scales.

There is an alternate definition of curvature, which calculates how fast the unit normal length decreases. Then, one has $\nabla L = 2\mathbf{N}\sin(\theta/2)$. This definition agrees in the continuous case, but doesn't preserve Theorem 3.2. In the limit, since $\sin\theta \approx \theta$, they're the same, but that's not what we are worried about.

The discrete curvature does converge in the limit, though this requires knowing a lot about the conditions: types of curves, types of convergence, where the points are sample, etc. Once again, there is differing iscrete behavior, but the same convergence in the end.

Next, step into the third dimension. There are very few papers about discretizing the Frenet frame, but one from applied physics, called *Structure Determination of Membrane Proteins Using Discrete Frenet Frames and Solid State NMR Restrictions*, by Achuthan and Quine, gives a suggestion. Here, the tangent can be approximated from a set of points $\mathbf{p}_1, \ldots, \mathbf{p}_n$, so that $\mathbf{t}_j = (\mathbf{p}_{j+1} - \mathbf{p}_j)/\|\mathbf{p}_{j-1} - \mathbf{p}_j\|$. Then, $\mathbf{b}_j = \mathbf{t}_{j-1} \times \mathbf{t}_j$, and the cross product game can be slightly modified to $\mathbf{n}_j = \mathbf{b}_j \times \mathbf{t}_j$.

Then, going from one frame to the next can be thought of as applying a matrix of rotation: $\mathbf{t}_k = R(\mathbf{b}_k, \theta_k)\mathbf{t}_{k-1}$, etc, and the derivative converges to the curvature. In the unit binormal case, one obtains the torsion. This is useful in the detection of kinked alpha helcies in proteins. This matrix contruction converges in the limit, but amazingly doeesn't require the curve to be smooth!

[4]This is a preview of something called DEC, which will be seen in a few weeks. Here, the point is a zero-dimensional cell, and a one-dimensional quantity is calculated.

However, just looking at the segments isn't always enough, as in the case of discrete elastic rods. For example, if someone twists and then compresses a shoelace, it bunches in ways that are difficult to simulate. This has applications to, for example, modelling hair.[5] This doesn't work in the piecewise linear case.

Thus, rather than just viewing the object as a segment, it might be helpful to also represent the twist in some way. An adapted framed curve is used for this: at each point one has two vectors $\mathbf{m}_1$ and $\mathbf{m}_2$, which encode the normal vector as it twists. These vectors are called the material frame. In order to simulate these curves, one could write down a potential energy called the bending energy:

$$E_{\text{bend}}(\Gamma) = \frac{1}{2}\int_\Gamma \alpha\kappa^2 \,\mathrm{d}s.$$

This is a form of potential energy that resists curvature, so to speak, where α is some constant. In the discrete case, the integral can be approximate by a sum, and in general can be computed in a closed form. Additionally, the curvature normal $\kappa\mathbf{n} = \mathbf{t}'$ can be written as

$$\begin{aligned}\kappa\mathbf{n} = \mathbf{t}' &= (\mathbf{t}'\cdot t)\mathbf{t} + (\mathbf{t}'\cdot\mathbf{m}_1)\mathbf{m}_1 + (\mathbf{t}'\cdot\mathbf{m}_2)\mathbf{m}_2\\ &= (\mathbf{t}'\cdot\mathbf{m}_1)\mathbf{m}_1 + (\mathbf{t}'\cdot\mathbf{m}_2)\mathbf{m}_2\\ &= \omega_1\mathbf{m}_1 + \omega_2\mathbf{m}_2,\end{aligned}$$

where ω_1, ω_2 are defined by the above. The above calculation depended on $(\mathbf{t}, \mathbf{m}_1, \mathbf{m}_2)$ forming an orthonormal basis of $\mathbb{R}^3$. Then, there is another potential energy called the twisting energy:

$$E_{\text{twist}}(\Gamma) = \frac{1}{2}\int_\Gamma \beta m^2 \,\mathrm{d}s,$$

where $m = \mathbf{m}_1'\cdot\mathbf{m}_2 = -\mathbf{m}_1\cdot\mathbf{m}_2'$ (this latter equality can be verified by the chain rule and the fact that lots of dot products cancel each other out.)

Next question: what basis should these be written in? The standard basis is easier to think about, but there are more frames than just the Frenet frame that have nice properties.[6] In particular, there is a frame called the Bishop frame made by choosing the most relaxed frame, so to speak. First, choose an orthonormal basis at the start of the curve. Then, the rest of the frame is given by solving a differential equation: if $\Omega = \kappa\mathbf{b}$, called the Darboux vector, then $\mathbf{t}' = \Omega\times\mathbf{t}$, $\mathbf{u}' = \Omega\times\mathbf{u}$, and $\mathbf{v}' = \Omega\times\mathbf{v}$ gives the whole frame $(\mathbf{t}, \mathbf{u}, \mathbf{v})$ (where $\mathbf{t}$ is again the tangent vector). Then, one has $\mathbf{m} - 1 = \mathbf{u}\cos\theta + \mathbf{v}\sin\theta$ and $\mathbf{m}_2 = -\mathbf{u}\sin\theta + \mathbf{v}\cos\theta$, making simulation very easy: the twisting energy is just

$$E_{\text{twist}}(\Gamma) = \frac{1}{2}\int_\Gamma \beta(\theta')^2 \,\mathrm{d}s.$$

Then, looking at a structure called a discrete Kirchoff rod, one can easily obtain the i^{th} tangent vector as $\mathbf{t}^i = \mathbf{e}^i/\|\mathbf{e}^i\|$ (where $\mathbf{e}^i$ is the i^{th} segment in the piecewise linear curve; subscripts indicate vertices and superscripts indicate edges). Then, one has *another* definition of curvature $\kappa_i = 2\tan(\phi_i/2)$, where ϕ_i is the angle between $\mathbf{e}^i$ and $\mathbf{e}^{i+1}$. After a bunch of algebra, one also has

$$(\kappa\mathbf{b})_i = \frac{2\mathbf{e}^{i-1}\times\mathbf{e}^i}{\|\mathbf{e}^{i-1}\|\|\mathbf{e}^i\| + \mathbf{e}^{i-1}\cdot\mathbf{e}^i}.$$

Then, the bending energy can be approximated as

$$E_{\text{bend}}(\Gamma) = \frac{\alpha}{2}\sum_i \left(\frac{(\kappa\mathbf{b})_i}{\ell_i/2}\right)^2 \frac{\ell_i}{2} \approx \alpha\sum_i \frac{\|(\kappa\mathbf{b})_i\|^2}{\ell_i},$$

which can be converted into the pointwise form and integrated.

Finally, one can talk about parallel transport. This is actually easyier to understand in the discrete case than in the continuous one. The goal is to find an orthogonal operator P such that $P_i(\mathbf{t}^{i-1}) = \mathbf{t}^i$ that preserves the binormal and orthogonality: $P_i(\mathbf{t}^{i-1}\times\mathbf{t}^i) = \mathbf{t}^{i-1}\times\mathbf{t}^i$. This is enough to completely specify it; thus, $\mathbf{u}^i = P_i(\mathbf{u}^{i-1})$ and $\mathbf{v}^i = \mathbf{t}^i\times\mathbf{u}^i$. Then, the material frame can be calculated explicitly by the formulae given above, and the twisting energy is just

$$E_{\text{twist}}(\Gamma) = \beta\sum_i \frac{(\theta^i - \theta^{i-1})^2}{\ell_i},$$

where θ^0 can be arbitrary. This allows for some reasonable simulations of physical processes, almost in real-time.[7]

---

[5] See *Discrete Elastic Rods*, Bergou, Wardetzky, Robinson, Audoly, and Grinspun, *SIGGRAPH* 2008.

[6] Richard L. Bishop, *There Is More Than One Way to Frame a Curve.*

[7] See the paper *Discrete Viscous Threads*, by Miklós Bergon, Basile Audoly, Etienne Vouga, Max Wardetzky, and Etian Grinspun, which announced the "first numerical fluid-mechanical sewing machine."

The takeaway is that there is one discrete (piecewise-linear) curve, but three ways to get the curvature: the angle θ, the gradient of the arc length $2\sin(\theta/2)$, and the parallel-transport formula $2\tan(\theta/2)$. In the limit these agree, but in the discrete case an intelligent decision must be used. Addiitonally, the best structures in classical differential geometry (e.g. the Frenet frame) that are easier to apply theoretically are very hard to apply. Half of the problem is just writing down the proper coordinates and degrees of freedom.

4. The Definition of a Surface: 4/10/13

"Good morning, guys... and good morning to everybody else who is still sound asleep."

The first concept necessary to define surfaces is the notion of the tangent space. Euclidean space also has a tangent space; the tangent space of $\mathbb{R}^n$ at a point p is denoted $T_p\mathbb{R}^n$. The idea here is to pick a point p and then choose three linearly independent vectors $E_{1,p}$, $E_{2,p}$, and $E_{3,p}$ that originate at p. This can be just the standard basis again, but it can also be a moving frame which depends on p. This is a fairly boring tangent space, in that it's just a replica of $\mathbb{R}^3$ in some other basis.

A standard construction of a tangent vector (i.e. an element of the tangent space) $X_p \in T_p\mathbb{R}^3$, there exists a curve $c : I \to \mathbb{R}^n$ such that $c(0) = p$ and $\dot{c}(0) = X_p$. Thus, one says that $X_p = \frac{\mathrm{d}c}{\mathrm{d}t}\big|_{t=0}$, where $c : [-\varepsilon, \varepsilon] \to \mathbb{R}^3$ is some small curve, and $c(0) = p$. This curve is not unique; examples include the straight line $c(t) = tX_p + p$, but then it is possible to bend it in some way, but at a positive distance from p, to get some other curve.

An important aspect of the tangent space is that the origin is moved to p. Thus, even though the space is basically the same as $\mathbb{R}^3$, the fact that the vectors originate at p rather than 0 means they aren't related to those originating at the origin, and that they can't be added or such without some explicit translation.

The union of all of the tangent spaces in $\mathbb{R}^3$ is called the tangent bundle: $T\mathbb{R}^3 = \{T_p\mathbb{R}^3 \mid p \in \mathbb{R}^3\}$. This is a six-dimensional object: there are three ways to pick p, and three dimensions within each $T_p\mathbb{R}^3$.

Definition. Suppose $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable. Denote coordinates in $\mathbb{R}^n$ as $(x^1, \dots, x^n)$ and the component functions of f as $f^1, \dots, f^m$, so that $f(x^1, \dots, x^n) = (f^1(x^1, \dots, x^n), \dots, f^m(x^1, \dots, x^n))$. Then, there are nm partial derivatives, and they can be placed into a matrix

$$Df_p = \begin{pmatrix} \frac{\partial f^1}{\partial x^1}(p) & \cdots & \frac{\partial f^1}{\partial x^n}(p) \\ \vdots & \ddots & \vdots \\ \frac{\partial f^m}{\partial x^1}(p) & \cdots & \frac{\partial f^m}{\partial x^n}(p) \end{pmatrix} \in \mathbb{R}^{m\times n}.$$

This is a linear transformation $Df_p : T_p\mathbb{R}^n \to T_{f(p)}\mathbb{R}^m$: f sends p to $f(p)$, and the differential maps the tangent space at the source point to the tangent space at the target point. This can be seen by using the curve representation of vectors in the tangent space: let $c : [\varepsilon, \varepsilon] \to \mathbb{R}^n$ such that $c(0) = p$ and $\frac{\mathrm{d}c}{\mathrm{d}t}(0) = X_p = (a_1, \dots, a_n)$ in the standard basis. Then, the point p can be moved forward under f, but so can the entire curve c, to get the curve $f(c(t))$, or the conposition $f \circ c$, which is a curve $[-\varepsilon, \varepsilon] \to \mathbb{R}^m$. Then, the tangent vector at $f(p)$ is $\frac{\mathrm{d}}{\mathrm{d}t}(f \circ c(t))\big|_{t=0} \in T_{f(p)}\mathbb{R}^m$. This can be related to X_p using the Chain Rule:[8]

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} f \circ c(t)\Big|_{t=0} &= \frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} f^1(c(t)) \\ \vdots \\ f^m(c(t)) \end{pmatrix} = \begin{pmatrix} \frac{\mathrm{d}}{\mathrm{d}t} f^1(c(t))\big|_{t=0} \\ \vdots \\ \frac{\mathrm{d}}{\mathrm{d}t} f^m(c(t))\big|_{t=0} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n \frac{\partial f^1}{\partial x^i}(c(t))\ \frac{\mathrm{d}c^i}{\mathrm{d}t}(t)\Big|_{t=0} \\ \vdots \\ \sum_{i=1}^n \frac{\partial f^m}{\partial x^i}(c(t))\ \frac{\mathrm{d}c^i}{\mathrm{d}t}(t)\Big|_{t=0} \end{pmatrix} \\ &= \begin{pmatrix} \sum_{i=1}^n \frac{\partial f^1}{\partial x^i}(p)a^i \\ \vdots \\ \sum_{i=1}^n \frac{\partial f^m}{\partial x^i}(p)a^i \end{pmatrix} = \begin{pmatrix} \frac{\partial f^1}{\partial x^1}(p) & \cdots & \frac{\partial f^1}{\partial x^n}(p) \\ \vdots & \ddots & \vdots \\ \frac{\partial f^m}{\partial x^1}(p) & \cdots & \frac{\partial f^m}{\partial x^n}(p) \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = Df_p \cdot X_p. \end{aligned}$$

Recognizing the matrix multiplication here is helpful in understanding how this is computed.

The differential says a lot about the function f. In differential geometry, the mapping properties of the differential are important: whether it is injective or surjective, what its kernel, image, eigenvalues, etc., and all of these say things about f. For example, the differential has a rank as a matrix (i.e. the dimension of the image).

Definition. A function f has locally constant rank K on an open set $\Omega \subseteq \mathbb{R}^n$ if $\operatorname{rank}(Df_p) = K$ for all $p \in \Omega$.

Then, one has the following results, which indicate how Df_p affects f near p, provided f has locally constant rank.

[8] It seems that every important calculation in differential geometry involves either the Product Rule or the Chain Rule.

Theorem 4.1 (The Rank Theorems)**.** *Suppose $f : \mathbb{R}^n \to \mathbb{R}^m$ has locally constant rank on a set $\Omega \subseteq \mathbb{R}^n$. Then, there exists a change of coordinates φ on $\mathbb{R}^n$ and ψ on $\mathbb{R}^m$ such that the map $\widetilde{f} = \psi \circ f \circ \varphi^{-1}$ (i.e. f written in the new coordinates) has the following properties:*

Case 1. If Df_p is injective for all $p \in \Omega$, then $n \leq m$ and $\operatorname{Ker}(Df_p) = \{0\}$. Then, $\widetilde{f}(x^1, \ldots, x^n) = (x^1, \ldots, x^n, 0, \ldots, 0)$.

Case 2. If Df_p is surjective for all $p \in \Omega$, then $n \geq m$ and $\widetilde{f}(x^1, \ldots, x^m, x^{m+1}, \ldots, x^n) = (x^1, \ldots, x^m)$; thus, it is a projection.

Case 3. If Df_p is bijective for all $p \in \Omega$, then $n = m$ and $\widetilde{f}(x^1, \ldots, x^n) = (x^1, \ldots, x^n)$. Thus, both f and $\widetilde{f}$ are invertible, since φ and ψ are.

Case 4. Suppose Df_p has rank k for all $p \in \Omega$ (i.e. neither injective nor surjective), then $k \leq \min(m, n)$ and $\widetilde{f}(x^1, \ldots, x^n) = (x^1, \ldots, x^k, 0, \ldots, 0)$.

This is a rectification (straightening-out) theorem. The intuition behind the coordinate change is that curves such as $x^2 + y^2 = c$ can be changed into curves such as $r^2 = c$ through the standard polar transform. This theorem promises a similar result in general. However, these are local coordinates, depending on Ω. An example of Case 1 is a curve: a straight line maps onto a curve, but after the coordinate change, the line is mapped to a straight line.

This entire theorem can be summarized as: *if Df_p has constant rank, then f behaves like Df_p near p.* Strictly speaking, proving this relies on two key theorems from multivariate analysis:

Theorem 4.2 (Inverse Function Theorem)**.** *If $f : \mathbb{R}^n \to \mathbb{R}^n$ is smooth and Df_p is bijective for some $p \in \mathbb{R}^n$, then f is invertible on some neighbrhood of p.*

Checking that Df_p is bijective is straightforward, since it is true iff $\det(Df_p) \neq 0$.

Theorem 4.3 (Implicit Function Theorem)**.** *Suppose $F : \mathbb{R}^k \times \mathbb{R}^n \to \mathbb{R}^n$ is smooth, $D_2F_{(p,q)}$ is bijective, and $F(p, q) = 0$. Then, the equation $F(x, y) = 0$ can be solved for points (x, y) near (p, q) in the sense that there exists a $g : \mathbb{R}^k \to \mathbb{R}^n$ defined near q such that $q = g(p)$ and $F(x, g(x)) = 0$.*

This can be thought of as solving a system of n equations in $n + k$ unknowns: there will be k free variables, and the dependencies among the rest will be given by $(x, g(x))$. In the above theorem, D_2F is the $n \times n$ derivative matrix in only the second coordinate, not a second erivative.

Example 4.1. Suppose $n = 1$ and $k = 2$, and $F(x, y, z) = x^2 + y^2 + z^2 - 1$. Then, $F(x, y, z) = 0$ when $(x, y, z) \in \mathbb{S}^2$ (i.e. the unit sphere in 2 dimensions), so $\mathbb{S}^2$ is the set of solutions of the equations. Solving for z, one gets $z = \pm\sqrt{1 - x^2 - y^2}$. Near $(0, 0, 1)$, the positive value is taken, and the solution is $g(x, y) = \sqrt{1 - x^2 - y^2}$, giving coordinates $(x, y, \sqrt{1 - x^2 - y^2})$.

Bringing in the Implicit Function Theorem, $D_2F = \frac{\partial F}{\partial z} = 2z$, so the equation can be solved locally provided $z \neq 0$. This makes, sense, because $z = 0$ is precisely where the formula given above by g fails: then, the derivative is badly defined, and the sphere has a vertical tangent. But everywhere else, it's absolutely fine.

In $\mathbb{R}^3$ in particular, there are three useful representations of surfaces, all of which have appeared somewhere else in these notes.

- Graphs of functions $f : \mathbb{R}^2 \to \mathbb{R}$. This has the important limitation that not every surface is a graph, such as the sphere.
- Level sets of functions $F : \mathbb{R}^3 \to \mathbb{R}$. Using the Implicit Function Theorem, it is possible to relate the level sets of one function as the graph of another: the space of solutions to $F(x, y, z) = 0$ is two-dimensional, so that it is a surface.
- Parameteric surfaces $\sigma : U \to \mathbb{R}^3$, where U is open in $\mathbb{R}^2$ and $\sigma(u^1, u^2) = (\sigma^1(u^1, u^2), \sigma^2(u^1, u^2), \sigma^3(u^1, u^2))$. This has the useful relation with level sets in that $F(\sigma(u))$ is constant. These surfaces are the generalization of curves, with $\mathbb{R}^2$ as the parameter domain rather than $\mathbb{R}$. This requires $D\sigma_p$ to be injective for all $p \in U$.

The sphere in polar (spherical) coordinates is an excellent example of a parametric surface: if θ is the angle of a point from the vertical (azimuth) and φ is the longitude, then the sphere is parameterized by $\sigma(\theta, \varphi) = (\cos\varphi \sin\theta, \sin\varphi \sin\theta, \cos\theta)$. Here, the parameter domain is $[0, \pi) \times [0, 2\pi)$ (not closed so that it remains injective).

5. Surface Geometry: 4/15/13

There are two formal definitions of a surface in differential geoemtry. This class will consider the less tricky one (i.e. not the general idea of an abstract manifold).

The key idea here is that graphs, level sets, and parameterizations don't cover all surfaces: not all surfaces are graphs, and finding the function that gives level sets is often difficult: it's a set of solutions of a system of equations, which can be difficult to solve. However, if it is solvable, it offers a compact representation of a huge family of surfaces.

Finally, for parameterizations, only part of some surfaces might be parameterizable at one time, and one has to keep track of these local coordinates and how they overlap. All of these methods have the additional issue that none of their representations are unique. This is both a theoretical and a practical problem, and leads to the following solution:

Definition. A set $S \subset R^3$ is a regular surface if for each $p \in S$ there exists an open neighborhood $V \subseteq \mathbb{R}^3$ of p, and open neighborhood $U \subseteq \mathbb{R}^2$, and a parameterization $\sigma : U \to V \cap S$ such that:

(1) $\sigma = (\sigma^1, \sigma^2, \sigma^3)$, such that each σ^i is differentiable.[9]
(2) σ is invertible as a map from U onto $V \cap S$ and has a continuous inverse. This is a nuance, because a map from $\mathbb{R}^2 \to \mathbb{R}^3$ cannot be invertible. However, restricted to these sets, it is possible to choose such a σ.
(3) $D\sigma_q$ is injective for all q (i.e. $\det((D\sigma_q)^{\mathrm{T}} D\sigma_q) \neq 0$)., so that σ preserves linear independence.

A regular surface is the analogue to a regular curve (i.e. one where the tangent vector never vanishes). The geometrical intuition is that σ covers a subset of the surface as a map from the parameter domain.

Example 5.1. Graphs are surfaces in this definition. If $f : \mathbb{R}^2 \to \mathbb{R}$, then one can take the surface $S = \{(x, y, f(x, y)) \mid x, y \in \mathbb{R}\}$. However, this comes with the caveat that the vertical line test must still hold and there cannot be vertical tangent lines.

The standard form for the graphical expression (i.e. the definition of S above) gives $\sigma(x, y) = (x, y, f(x, y))$. This satisfies all of the axioms, so graphs are in fact surfaces. Looking particularly at the third axiom, which might be the least clear,

$$D\sigma_{(x,y)} = \begin{pmatrix} \frac{\partial \sigma^1}{\partial x} & \frac{\partial \sigma^1}{\partial y} \\ \frac{\partial \sigma^2}{\partial x} & \frac{\partial \sigma^2}{\partial y} \\ \frac{\partial \sigma^3}{\partial x} & \frac{\partial \sigma^3}{\partial y} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix}$$

$$D\sigma^{\mathrm{T}} D\sigma = \begin{pmatrix} 1 & 0 & \frac{\partial f}{\partial x} \\ 0 & 1 & \frac{\partial f}{\partial y} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} 1 + \left(\frac{\partial f}{\partial x}\right)^2 & \frac{\partial f}{\partial x}\frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial x}\frac{\partial f}{\partial y} & 1 + \left(\frac{\partial f}{\partial y}\right)^2 \end{pmatrix}$$

Thus, the determinant of $D\sigma^{\mathrm{T}} D\sigma$ can be calculated to be $1 + (\frac{\partial f}{\partial x})^2 + (\frac{\partial f}{\partial y})^2$, which can never be zero. Note that this requires f to be smooth, so non-smooth functions may have singular points on their surfaces (e.g. a cone).

Example 5.2. A sphere is a regular surface, as it can be written as the union of six parameterizations: take $x^2 + y^2 + z^2$, solve for one of x, y, or z, and take the positive and negative solutions in each case. These are all necessary because the equator causes issues.

Another example is the inverse image of a regular value, which will be seen in a bit.

Now that the geometry is established, some differential things will be needed as well. A curve in a surface $\gamma : I \to \mathbb{R}^3$ is always contained in the surface: $\gamma(t) \in S$ for all $t \in I$. Thus, $\gamma(t) = \sigma(\gamma_0(t))$, and such a γ_0 can always be found locally. For example, if $\gamma(t) = (t, t, \sqrt{1 - 2t^2})$ on the unit sphere, then $\gamma_0(t) = (t, t)$.

One also has the notion of coordinate curves: if U has local coordinates u^1 and u^2, then there are curves that are parallel to the u^1- or u^2-axes: $\gamma = (t, k)$ or $\gamma = (k, t)$ for some conatant k. These curves map into sime network of curves on the surface. This means one can visually represent the injectivity of $D\sigma$: two vectors pointing in the u^1- and u^2-directions and the square they generated are mapped to linearly independent vectors, and the square doesn't collapse.

Thus, one can take the tangent vectors $\dot{\gamma}_0(u)$ of a curve $\gamma_0 : I \to U$ and push them all to the surface as tangent vectors v of $\sigma \circ \gamma_0$, given by $v = D\sigma_u \cdot \dot{\gamma}_0(u)$. Importantly, v is tangent to the surface.

Definition. Let $\sigma : U \subseteq \mathbb{R}^2 \to V \cap S \subseteq \mathbb{R}^3$ be a parameterization of a subset of a surface S and choose a $p \in S$ such that $p = \sigma(u)$ for some $u \in U$. Then, the tangent plane is defined as $T_p S = \mathrm{Im}(D\sigma_u) \subseteq T_{\sigma(u)}\mathbb{R}^3$.

This definition means that the tangent plane is the push-forward of all possible tangent vectors of curves in the parameter domain through p.

This illustrates a more general principle of differential geometry: the concept must be established with some parameterization, but then can be shown to be independent of parameter. The general idea of the proof is if $\sigma : U \to \mathbb{R}^3$ and $\tau : U' \to \mathbb{R}^3$ are two different parameterizations of the same section of the surface, $\sigma \circ \tau^{-1} : U' \to U$ is a smooth bijection. Then, $\mathrm{Im}(D\sigma_u) = \mathrm{Im}(D(\sigma \circ \tau^{-1} \circ \tau)_u) = \mathrm{Im}((D\sigma \circ \tau^{-1})_{\tau(u)} D\tau_u) = \mathrm{Im}(D\tau_u)$, since the differentials can be separated by the chain rule and mutiplying by an invertible matrix doesn't change the column space of a matrix.

[9]Sometimes, other conditions are specified, such as smoothness (equivalently, being C^∞ or infinitely differentiable.)

Thus, the tangent space is a geometric object, but finding a basis isn't independent of parameterization. Given a parameterization $\sigma : U \to \mathbb{R}^3$, though, one can take $v_1 = D\sigma_u \binom{1}{0} = \frac{\partial \sigma}{\partial u^1}$ and $v_2 = D\sigma_u \binom{0}{1} = \frac{\partial \sigma}{\partial u^2}$, so (v_1, v_2) is the differential.

Example 5.3. For a graph given by $f(x, y)$, $v_1 = \begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial x} \end{pmatrix}$ and $v_2 = \begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial y} \end{pmatrix}$.

6. Discrete Surfaces: 4/17/13

The basic assumption today is that all surfaces are manifold meshes: each edge is incident on one or two faces, and each vertex forms a fan (so there are no bowtie-like structures of triangles meeting only at a point). This assumption, which underlines many discrete surface algorithms, is already a huge assumption, and doesn't always happen in real life.

By the Mean Value Theorem and Taylor's Theorem, these are good approximations in the linear case: the error is quadratic, and there exists a point on the approximation where the tangent line is the same. There are other advantages of triangle meshes: they are easy to render, and they can form lots of different topologies (e.g. spheres with n handles). Finally, subdivision and refinement of meshes make life easier. Thus, triangle meshes encode both the geometry and topology of a surface.

The geometry of a triangle mesh is fairly clear: where are the vertices? Topological statements might include the number of holes in the surface, whether two vertices are connected, etc. This is a higher-level property than some of the geometric ones.

Formally, a triangle mesh is a set $V = \{v_1, \dots, v_n\} \subset \mathbb{R}^n$ of vertices, a set $E = \{e_1 \dots, e_n\} \subset V \times V$ of edges, and a set $F = \{f_1, \dots, f_n\} \subset V \times V \times V$ of faces, where the edges, faces, and vertices are all adjacent and satisfy the manifold conditions. This can be generalized (e.g. tetrahedra; the more general term is simplicial complex).

Definition. The valency or degree of a vertex in a triangle mesh is the number of edges coming out of it.

Definition. The Euler characteristic of a surface with V vertices, E edges, and F faces is $\chi = V - E + F$.

If a surface has g holes, then its characteristic is $\chi = 2 - 2g$.

Applying this to a triangle mesh, each edge is adjacent to two faces, and each face has three edges, so $2E = 3F$, and $V - F/2 = \chi$ by substitution. Additionally, if a surface has few or no holes (which is a reasonable assumption in applications), then χ is small, so $V - F/2$ is small, and $F \approx 2V$. Of course, thee are lots of counterexamples, but in these cases, $E \approx 3V$ and $F \approx 2V$, so the average valency of the vertices is about 6: there are three times as many vertices as edges, but when summing up the degrees, each edge is double-counted.

Another useful piece of information about a triangle mesh is orientability. Something like a Möbius strip is nonorientable. It's still a surface, but some algorithms or definitions require orientability. Formally, a surface is orientable if there exists a continuous vector field everywhere normal to the surface. Of course, on a triangle mesh, continuity is suspect, so this has to be sidestepped: each face is given an orientation (clockwise or counterclockwise), and the mesh is orientable if at every edge, the two adjacent faces are going in opposite directions (at that edge; not clockwise or counterclockwise over the whole face). This relates to the continuous definition, since using the right-hand rule for cross products, one can obtain the normal vector field, though its continuity can't be really talked about. It will be assumed that meshes are orientable for much of this class, and it is worth checking whether various algorithms require orientability.

How should a triangle mesh be represented? The simplest format is known as "triangle soup:" the three vertices of each face are stored on a line. Vertices that are shared between faces are repeated. This encodes absolutely no topological information, but is incredibly easy to render.

This can be improved by keeping a numbered list of vertices and have faces refer to the vertex numbers. This is the decision that the `.obj` format makes.

Consider the following algorithm:

```
for i=1 to n
    for each vertex v
        v = .5*v + .5*(average of neighbors);
```

The `.obj` format has no nice oncept of neighbors, so this will not work quickly. But it looks like it should be an easy operation, so maybe there's a better data structure. Other localized operations include finding neighboring faces to an edge, or edges adjacent to a face or vertex, etc.

One of the more popular solutions to this is the half-edge, in which the vertices and faces are stored as before, and a list of half-edges are also stored. This is specifically tuned to orientable polygonal meshes. The difference between edges and half-edges is that half-edges are oriented, so each edge corresponds to two half-edges with opposite

directions. Each half-edge are associated with a single face. Then, each vertex only needs to store its outgoing half-edge.[10] Each face stores an arbitrary half-edge. Then, each half-edge stores its flip (the other half-edge in the same place), the next half-edge on the same face (pointed two by the orientation), the face associated with the half-edge, and the base vertex. This is a topological data structure.

Returning to iterating over the neighbors, it's suddenly much easier. Given some vertex v, the neighboring vertices are the vertices pointed to by the half-edges leaving the vertex. Since the half-edge doesn't store the vertex it points to, use the flip, which does store it. Then, since the flip is on a different face, following it to the next half-edge gives another vertex. This method does weird things on the boundary, and the conventional solution is to tag boundary edges and handle them differently.

There are lots of other representations of triangle meshes, and the one to use depends on what application or algorithm is the most important.

Looking at this more topologically, a face is of dimension 2, an edge is of dimension 1, and a vertex is of dimension 0. Thus, one can think of the boundary operator ∂ that sends a face to an edge and an edge to a vertex. This operator can track orientation, making it in some sense a signed operator. Addiitonally, $\partial^2 = 0$: the orientations all cancel out.

Turnign to scalar functions: a function from a surface to $\mathbb{R}$ can be thought of as a heat-map on the surface. The most straightforward representation is to pick it on the vertices: $f \in \mathbb{R}^{|V|}$. However, this doesn't lead to a straightforward definition of the integral of f.

One way to approch this is the finite elements standpoint, which uses something called a hat function, which approixmates the vertex representation of f locally, but is continuous. Then, one could sum all of the hat functions to obtain a reasonable continuous approximation. Alternatively, one could use dual cells, which are faces corresponding to the vertices. Then, the function can be integrated on the dual faces. Note that the dual of a triangle mesh is a polygon mesh, but not always a triangle mesh. The valency in the dual mesh is always 3 (except on the boundary), since each face in the original mesh is a triangle, so there are three neighboring faces. These dual structures both

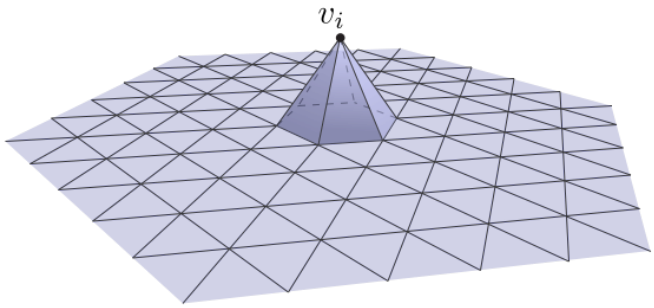

FIGURE 2. A hat function on a discrete surface. Source

represent the same surface, and can be related. The quad-edge data structure encodes this, where one can rotate from one half-edge to its dual. This also supports operations such as vertex removal or edge or face collapse. The general conclusion is that more complicated data structures enable simple traversal but require a lot more bookkeeping.

There are lots of other ways to store surfaces. For example, in a point cloud, it may not be possinble to get a reasonable triangle mesh. Here, a function of distance to the nearest point is used, and makes some operations (e.g. subtraction) much easier. Alternatively, one could use smooth-particle hydrodynamics, where points are represented as spheres. Given this point representation, extracting a triangle mesh is a meaningful question. Interestingly, representing things such as surface tension are best expressed with differential-geometric quantities such as curvature.

Finally, it's worth asing how this geometry may be obtained. Probably the cleanest way is design software (e.g. in a CAD model, where the most important property is that it looks right). Then, some sort of subdivision operator can be used to obtain a nicer representation. In the case of medical imaging, volumetric CT scanners are used, so one can use a method called marching cubes to extract a function based on the values at the vertices.[11] This does involve a lot of cleanup, however. Given a point cloud, one could make a signed distance function to determine where the inside and outside of an object are. Then, the surface is just the zero set. However, one could also use this to obtain the tangent space near a point, and thus patch the tangent spaces at various points together. Finally, one can use a method called Delaunay triangulation (dual to Voronoi cells) gives a triangulation with a well-behaved dual mesh as well. A more general application to point clouds is Poisson representation, which solves a PDE on the volume around a point set to obtain the normals. This creates a function whose level sets include the surface.

[10]Some implementations store more information than the structure mentioned here, but that's not all that important

[11]Apparently this is patented. I guess we've really wandered away from mathematics now.

7. Extrinsic Curvature: 4/22/13

This class will begin dealing with Riemannian geometry, rather than just differential geometry.

The line through the normal is a geometric object, since it's determined by the tangent plane. However, the normal line (unit normal) might not be — it can depend on orientation. Additionally, there are non-orientable surfaces such as the Möbius strip, which doesn't admit a consistent set of normal directions.

In order to actually calculate the normal, consider the case where the surface is parameterized (which is always true locally). Suppose $T_pS = \operatorname{span}\{E_1, E_2\}$; then, $N = \frac{E_1 \times E_2}{\|E_1 \times E_2\|}$, so that N is always perpendicular to the tangent plane.

If the surface is foliated by level sets $F^{-1}(V) = \{x \in \mathbb{R}^3 \mid F(x) = v\}$, where $F : \mathbb{R}^3 \to \mathbb{R}$. Choose a level set $S = F^{-1}(v)$ for some v, and let $c : I \to \mathbb{R}^3$ be a curve contained in S such that $c(0) = p$ and $\dot{c}(0) = X \in T_pS$. Thus, $v = F(c(t))$ for all t, so

$$\begin{aligned} 0 = \frac{\mathrm{d}}{\mathrm{d}t} F(c(t))\bigg|_{t=0} &= \left\langle [DF_p]^{\mathrm{T}}, X \right\rangle \\ \implies N &= \frac{[DF_p]^{\mathrm{T}}}{\|DF_p\|}, \end{aligned}$$

so $N \perp T_pS$ again, since it was seen to be orthogonal. This assumes that v is a regular value of F, so that the surface is smooth there, and the derivative matrix is nonzero.

In Riemannian geometry, it is possible to determine surface integrals. If S is some surface and $f : S \to \mathbb{R}$, then the objective is to obtain some reasonable notion of $\int_S f$. The first idea is to work inside a parameterization. Let $\varphi : \mathcal{U} \to \mathbb{R}^3$ be a local parameterization of S. Thus, one has a function $f \circ \varphi : \mathcal{U} \to \mathbb{R}$, so it seems reasonable to define the integral as

$$\int_S f = \int_{\mathcal{U}} f \circ \varphi(u) \, \mathrm{d}u^1 \, \mathrm{d}u^2.$$

This is not a good definition, however, as it is highly dependent on parameterization. The chart map φ may also change the area, which leads to a better solution: φ takes small rectangles $R \subset \mathcal{U}$ and deforms them into parallelepipeds $\varphi(R)$. Then, one can take a Riemann sum:

$$\int_S f \approx \sum_i f \circ \varphi(u_i) \operatorname{Area}(\varphi(R_i)),$$

where the R_i are non-overlapping rectangles that cover $\mathcal{U}$ and $u_i \in R_i$ (the midpoints of the rectangles). Then, let $E_1 \wedge E_2$ represent the rectabgle spanned by E_1 and E_2, which has area $\|E_1 \times E_2\| = \sqrt{|\det(D\varphi_u^{\mathrm{T}} D\varphi_u)|}$. Thus, one can define the Riemannian area form $\mathrm{dArea} = \sqrt{|\det(D\varphi_u^{\mathrm{T}} D\varphi_u)|} \, \mathrm{d}u^1 \, \mathrm{d}u^2$, which gives

$$\int_S f \, \mathrm{dArea} = \lim \sum_i f(\varphi(u_i)) \sqrt{\det(D\varphi_u^{\mathrm{T}} D\varphi_u)}.$$

This is independent of parameterization, because the change-of-coordinate map gives determinants that cancel out, albeit after an involved calculation.

Let S be an orientable surface with unit normal at $p \in S$ given by N_p. Then, recall that the Gauss map of S is the mapping $n : S \to \mathbb{S}^2$ given by $n(p) = N_p$, where N_p is considered moved to the origin; since it has unit length, it is on the unit sphere. It happens that the Gauss map of a differentiable surface is also differentiable; its differential is $D_{n_p} : T_PS \to T_{N_p}\mathbb{S}^2$, called the shape map, and defines some interesting geometry. One can view S and $\mathbb{S}^2$ as embedded in the same $\mathbb{R}^3$, so T_pS and $T_{N_p}\mathbb{S}^2$ are parallel planes, so one can redefine $Tn_p : T_pS \to T_pS$. This is slightly problematic for abstract surfaces, but slightly different definitions still work out. Now, one can introduce the associated bilinear form:

Definition. The second fundamental form of S at p is the bilinear form $A_p : T_pS \times T_pS \to \mathbb{R}$ given by $A_p(V, W) = -\langle Dn_p(V), W \rangle$, with the sign by convention.

This measures the rate of change of the normal vector at p in the V-direction, projected onto W. Taking the inner product and matrix multiplication are linear operations, so this form is bilinear. The projection onto W is necessary to ensure independence of basis, and also makes everything a bit easier to deal with.

Proposition 7.1. *A_p is self-adjoint; that is, $-\langle Dn_p(V), W \rangle = \langle V, Dn_p(W) \rangle$, or $A_p(V, W) = A_p(W, V)$.*

Proof. Choose a basis: take $\varphi : \mathcal{U} \to S$, and let $E_1 = \frac{\partial \varphi}{\partial u^1}$ and $E_2 = \frac{\partial \varphi}{\partial u^2}$ be the basis. Then, a little magic happens, but it's of the same flavor as before: using the product rule,

$$[A_p]_{ij} = -\left\langle \frac{\partial N}{\partial u^i}, \frac{\partial \varphi}{\partial u^j} \right\rangle = -\frac{\partial}{\partial u^i} \left\langle N, \frac{\partial \varphi}{\partial u^j} \right\rangle + \left\langle N, \frac{\partial^2 \varphi}{\partial u^i \partial u^j} \right\rangle = \left\langle N, \frac{\partial^2 \varphi}{\partial u^i \partial u^j} \right\rangle,$$

since N and $\frac{\partial \varphi}{\partial u^i}$ are orthogonal. Now, since φ is well-behaved, its mixed partials are equal, so the above steps can be reversed:

$$= \left\langle N, \frac{\partial^2 \varphi}{\partial u^j \partial u^i} \right\rangle = -\left\langle \frac{\partial N}{\partial u^j}, \frac{\partial \varphi}{\partial u^i} \right\rangle = [A_p]_{ji}. \qquad \boxtimes$$

The second fundamental form has three interpretations; the first of these is extrinsic curvature. Let c be a curve in S with $c(0) = p$, and let $\mathbf{k}_c(0)$ be the geodesic curvature of c at zero. Then, $\dot{c}(0)$ is perpendicular to the surface, but $\ddot{c}(0)$ may or may not be parallel or perpendicular. However, once the inner product is taken with the normal, there is some information. In the arc-length parameterization,[12]

$$\begin{aligned}
\langle \mathbf{k}_c(0), N_p \rangle = \langle \ddot{c}(0), N_p \rangle &= \left\langle \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}c}{\mathrm{d}t}\right), N_p \right\rangle \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left\langle \frac{\mathrm{d}c}{\mathrm{d}t}, N_p \right\rangle - \left\langle \frac{\mathrm{d}c}{\mathrm{d}t}, \frac{\mathrm{d}N(c(t))}{\mathrm{d}t} \right\rangle \\
&= -\left\langle \frac{\mathrm{d}c}{\mathrm{d}t}, \frac{\mathrm{d}N(c(t))}{\mathrm{d}t} \right\rangle \\
&= -\langle \dot{c}(0), Dn_p(\dot{c}(0)) \rangle = A_p(\dot{c}(0), \dot{c}(0)).
\end{aligned}$$

Thus, the curvature is related to the second fundamental form, and in particular only depends on the geometry of Sat p. One can take the circle with the same curvture and $\ddot{c}$, which induces the same second fundamental form, and move it until it is tangent at S. The radius of this circle is called the radius of curvature, and is equal to $A(\dot{c}(0), \dot{c}(0))$.

Let V vary over all unit vectors in T_pS. Then, A_p takes on a minimum $k_{\min}$ and a maximum $k_{\max}$. Then, the mean curvature is defined to be $H = k_{\min} + k_{\max}$, and the Gauss curvature is $K = k_{\min} \cdot k_{\max}$. These are the trace and determinant, respectively, of A_p in an orthonormal basis. Additionally, the eigenvalues of A_p are called the principal curvatures of S at p (which are $k_{\min}$ and $k_{\max}$), and their corresponding eigenvectors are called the principal directions, which are orthogonal because A_p is a self-adjoint quadratic form. Here, the mean and Gauss curvatures are particularly nice because they can be computed without diagonalization, since the trace and determinant are invariant under choice of basis.

The second fundamental form is also related to the local shape of a surface, which will be discussed next week.

8. Computing Curvature: 4/24/13

"So we take two tangent vectors, V snd Dubya..."

Probably the most well-known facet of differential geometry is curvature. Unlike the previous lecture, this one will focus on computing the curvature of discrete surfaces.

Consider a surface S, a point $p \in S$, and its unit normal vector N. For some curve $\gamma \subset S$, the Gauss map $S \to \mathbb{S}^2$ measures the direction of the unit normal by translating it to the origin, so it can be identified with a point on the unit sphere.

Further recall the second fundamental $A_p(V, W) = -\langle DN_p(V), W \rangle$ for $V, W \in T_pS$. The differential DN_p can be thought of as a matrix or as an operator on tangent spaces. The second fundamental form can be used to obtain principal directions and curvatures: let the first principal direction for p be the direction in which the curvature is greater. Based on the eigenstructure of the fundamental form, we have the following theorem:

Theorem 8.1. *Let κ_1 be the direction of the greatest curvature and κ_2 be the direction of the smallest curvature. Then, they are always perpendicular and the curvature in any direction θ is $\kappa_1 \cos^2\theta + \kappa_2 \sin^2\theta$.*

The reason curvature is so important is because the curvature completely defines surface geometry: if the principal curves and directions are known, the surface itself can be reconstructed. The proof of this is slightly messy. This can be thought of as the analogue of Theorem 2.1. Intuitively, the surface is bowl-shaped if κ_1 and κ_2 have the same sign, and saddle-shaped if otherwise.

There are lots of different types of curvatures: the Gaussian curvature emphasizes differences such as that, and indicates whether the surface is elliptical or hyperbolic at a point. The mean curvature completely ignores these sorts of distinctions, but has other uses.

This can be used in applications: if one wishes to smoothen out a surface, this involves slightly editing the curvature, and then using it to reconstruct the surface. This ends up being a tricky problem. Additionaly, one could use curvature to mae a surface smoother at a point. Another application is to use curvature to highlight interesting geometric features. Finally, one can use curvature to re-mesh a surface. This method is used when the specific underlying mesh is different from the one that was obtained. Thus, one can use the curvature of one mesh to induce the next. However,

[12] Some of these steps can be condensed by he self-adjoint property.

this underlies a scary combinatorial question: it's not just a local problem, and the actual topology of the mesh has to be changed.[13]

Definition. Given a smooth surface S and a point $p \in S$, let the tangent plane be given by T_1 and T_2 in the principal directions. Then, let $T_\theta = T_1 \cos\theta + T_2 \sin\theta$. Then, the Taubin matrix is

$$M = \frac{1}{2\pi} \int_{-\pi}^{\pi} \kappa_\theta T_\theta T_\theta^{\mathrm{T}} \mathrm{d}\theta.$$

This is kind of useless if κ_1 and κ_2 are already known, but it has some nteresting properties.

Exercise 8.1. Show that the eigenvectors are N, T_1, and T_2, and that the eigenvalues are $3\kappa_1/8 + \kappa_2/8$ and $\kappa_1/8 + 3\kappa_2/8$.

The formula above is kind of magic; there doesn't seem to be an intuitive reason for the 8 to be there.

The Taubin matrix can be approximated by replacing the integral by a sum: denote adjacency by $\sim$. Then,

$$\widetilde{M}_{v_i} = \sum_{v_j \sim v_i} w_{ij} \kappa_{ij} T_{ij} T_{ij}^{\mathrm{T}}.$$

Here, w_{ij} is some choice of weights. This just weights the approximations of the unit normal. One reasonable model is that if v_i and v_j are on triangles with areas A_1 and A_2, $w_{ij} \propto A_1 + A_2$. This is a very fast algorithm, especially when combined with the half-edge data structure.

This has a problem, though: local estimates are very noisy. Sometimes, most of the surface looks fine and there are only a few localized issues. Smoothing the mesh to eliminate this is difficult, since it typically requires computing the curvature! This problem is quite difficult, and lots of papers have been written on it.

In some sense, the game can be described as the following:[14]

1. Collect local data for a surface.
2. ???
3. Profit!

Here, the second step is some tool in the differential-geometric toolbox, and the third involves a nice calculation of curvature.

The rest of the lecture contains engineering disguised as math. As with most applications, the approximation of curvature chosen should be motivated by the application (e.g. in animation, where it matters more that it's stable than it's so accurate).

One standard method for computing curvature can be found in Szymon Rusinkiewicz's paper *Estimating Curvatures and Their Derivatives on Triangle Meshes*. Note that each paper uses its own favorite notation, so some discretion is advised. This is probably because lots of these poeple took the undergraduate differential geometry courses but not graduate ones, which meant they didn't see the newest, more uniform notation.

Though mathematicians thing of the second fundamental form as an operator, in application it's helpful to think of it as a matrix:

$$\mathrm{II} = (D_u n \; D_v n) = \begin{pmatrix} \frac{\partial n}{\partial u} \cdot u & \frac{\partial n}{\partial v} \cdot u \\ \frac{\partial n}{\partial v} \cdot u & \frac{\partial n}{\partial v} \cdot v \end{pmatrix}.$$

Let $s = c_1 u + c_2 v$; then $\mathrm{II} \binom{c_1}{c_2} = D_s n$.[15]

Discretizing, consider a triangle with edgs e_0, e_1, e_2, such that e_0 is opposite the point with normal n_0, and similarly with n_1 and n_2. Then,

$$\mathrm{II} \begin{pmatrix} e_0 \cdot u \\ e_0 \cdot v \end{pmatrix} = \begin{pmatrix} (n_2 - n_1) \cdot u \\ (n_2 - n_1) \cdot v \end{pmatrix},$$

and similarly with the other two directions. In practice, this can be done by rotationg the tangent plane by the cross product, and then average the adjacent matrices together. Thus, this method is very stable: the system of matrix equations is overdetermined, so the least-squares solution is stable, and the average is taken on top of that.

A lot of papers play a similar game: take some theoretical formula, discretize it, and see how it works. Sometimes, this can involve some surprising areas of mathematics, including topology and Morse theory.

An alternate approach is to take some conserved property and build on that. There's way too much literature to easily condense it. Thus, the approch will be to take some theoretical quantity and try to preserve it, possibly at the expense of others.

---

[13] Apparently, the instructor implemented some of these methods in high school. Along the way his group somehow obtained a patent on Gaussian curvature...

[14] Yes, this was said in lecture.

[15] This notation is likely to be particularly confusing; today's $D_s n$ is usually thought of as $D_n s$: s is the vector that the differential is calculated from.

Theorem 8.2 (Gauss-Bonnet)**.** *Let M be a surface, and ∂ be the boundary operator. Then,*

$$\int_M K \,\mathrm{d}A + \int_{\partial M} k_g \,\mathrm{d}s = 2\pi\chi(M),$$

where χ is the Euler characteristc, K is the Gaussian curvatue, and k_g is the geodesic curvature (i.e. curvature projected onto the tangent plane).

This is statement of a global, topological property in terms of local, geometric ones, which is pretty impressive. However, its proof requires some more machinery than has been developed thus far, and will be deferred.

For polygonal cells, this simplifies: suppose that a triangle fan V at a vertex has a dual whose exterior angles are ε_j. Then, the geodesic curvature is just the sum of these angles, and

$$\int_V K \,\mathrm{d}A = 2\pi - \sum_j \varepsilon_j.$$

This reduces the problem of curvature of surfaces to that of curvature of curves,

Exercise 8.2. If V is a Voronoi cell, prove that additionally show that

$$\int_V K \,\mathrm{d}A = 2\pi - \sum_j \theta_j, \tag{1}$$

where the θ_j are the angles that the triangles make with the central vertex.

Note that the right-hand quantity is *not* curvature: it is integrated curvature, so for an approximation, one needs to divide by the area of V. However, it's easy to compute, and therefore

Definition. The Gaussian curvature of a mesh is defined to be the quantity that satisfies (1).

Using this, it is possible to prove the discrete analogue of the Gauss-Bonnet theorem:

$$\begin{aligned}\int_M K \,\mathrm{d}A &= \sum_i \int_{V_i} K \,\mathrm{d}A \\ &= \sum \left(2\pi - \sum_j \theta_{ij}\right) \\ &= 2\pi V - \sum_{i,j} \theta_{ij} \\ &= 2\pi V - \pi F = \pi(2V - F) = 2\pi\chi.\end{aligned}$$

where F is the number of triangles, since each triangle contributes an angle of $180°$.

Exercise 8.3. Show that the Gauss-Bonnet theorem does not hold for the Taubin curvature.

This definition of curvature can still be incredibly noisy, but it is consistent in precisely the way that was intended.

It also turns out that the curvature is related to maximizing surface area, just like arc length for curves. This is hard to compute for general smooth surfaces, but computing the surface area of a triangle mesh is stupidly easy. The surface area can be thought of as a function $A : R^{3|V|} \to \mathbb{R}$, where V is the set of vertices.

For a single triangle given by $\mathbf{e}$, $\mathbf{p}$, and $\mathbf{p}-\mathbf{e}$, let $\mathbf{n}$ be normal to $\mathbf{e}$ and $\mathbf{e}_\perp$ be binormal. Then, $\mathbf{p} = p_n\mathbf{p} + p_e\mathbf{e} + p_\perp\mathbf{e}_\perp$ (written in coordinate components) ans $A = b\sqrt{p_n^2 n + p_\perp^2}/2$. Thus,

$$\frac{\partial A}{\partial p_e} = 0, \qquad \frac{\partial A}{\partial p_n} = \frac{bp_n}{2\sqrt{p_n^2 + p_\perp^2}} = 0, \text{ and} \qquad \frac{\partial A}{\partial p_\perp} = \frac{bp_\perp}{2\sqrt{p_n^2 + p_\perp^2}},$$

so $\nabla_{\mathbf{p}} A = b\mathbf{e}_\perp/2$.

After an awful lot of high-school trigonometry and angle-chasing, one obtains a handy-dandy formula

$$\nabla_{\mathbf{p}} A = \frac{1}{2}((\mathbf{p}-\mathbf{r})\cot\alpha + (\mathbf{p}-\mathbf{q})\cot\beta,$$

where α is opposite the line $\mathbf{p}-\mathbf{r}$ and β is the angle opposite $\mathbf{p}-\mathbf{q}$. This formula will be very useful.

For a first application, one can obtain the gradient of the surface area by walking around a vertex. Then, its norm is the mean curvature. As points get closer to each other, this gets closer to zero.

Definition. The mean curvature normal integrated over a region V is given by

$$\int_V H\mathbf{n}\,\mathrm{d}A = \frac{1}{2}\sum_j (\cot\alpha_j + \cot\beta_j)(\mathbf{p}-\mathbf{q}_j),$$

where $\mathbf{p}$ is the vertex for a fan, $\mathbf{q}_j$ are the non-central vertices and have angles α_j and β_j.

This is probably less stable than the least-squares solution, but has other nice properties.

There are still lots of other ways to compute curvature: for applications in triangulation, one can use dihedral angles between two faces rather than the interior angles next to a vertex. Then, one can imagine smoothing out each triangle edge. Let β be the bending angle, which is consistent throughout the circular filling. Since it holds at every scale, it can be exploited to obtain curvature. Then, define the mean curvature to be satisfied by

$$\int_B H = \frac{1}{2}\beta\|\mathbf{e}\|,$$

where $\mathbf{e}$ is the edge in question. There may be sign questions, but they are small.

This can be used to retriangulate a surface; for example, if one wants to take two triangles that share an edge and flip the edge, this makes it very reasonable to calculate whether the flipped edge has higher or lower curvature.

Alternatively, one wants curvature in image processing: there are lots and lots of blur operators, so one wants definitions of curvature that are stable when blurred out. This is particularly useful when rendering, and can have surprisingly good accuracy by only combining curvatures where the surface has a similar geometry. This is what lies behind the bilateral filter. One could fit a smooth surface locally, or even use machine learning to create good approximations to some complicated curvature calculation. The upshot is that if you need some sort of curavture, try different approaches until one works.

9. Intrinsic Geometry: The Induced Metric and Geodesics : 4/29/13

First, a little more about the second fundamental form:

Example 9.1. Let S be the graph of a function $f:\mathbb{R}^2\to\mathbb{R}$. Without loss of generality, assume f vanishes to first order at the origin.[16] Then, in this standard form, the second fundamental form is really simple:

$$A_{(0,0)} = \left.\begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x\partial y} \\ \frac{\partial^2 f}{\partial x\partial y} & \frac{\partial^2 f}{\partial y^2}\end{pmatrix}\right|_{(0,0)}$$

This is only valid in these conditions, however. Since it's a quadratic form, one can look at its eigenvalues: thei eigendirections are orthogonal. Then, the surface can be locally characterized:

- If both eigenvalues have the same sign, the surface is elliptic (it looks like a bowl).
- If they have opposite signs, the surface is hyperbolic (saddle-shaped).
- If one of the eigenvalues is zero, it's a parabolic point.
- If both are zero, the surface is planar.
- If the eigenvalues are equal, the surface is called umbilic. This is important because the eigenbasis is not well-defined in this case.

This can be shown with Taylor's theorem; the second-order term is the first nonzero one and dominates the Taylor expansion, so the surface looks like the graph of a quadratic function. If one of the eigenvalues is zero, one has to go to the cubic, quartic, or higher terms, so the theorem is inconclusive, and a parabolic surface could have cubic or other such behavior.

Recall that the mean curvatureis the gradient of the surface area. The surface area can be thought of as a functional, so what does the gradient mean? Take some surface S and take some variation S_ε, which is a small change of the surface. Then, $\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\operatorname{Area}(S_\varepsilon)\big|_{\varepsilon=0}$ is the gradient. Thus, the area of the surface decreases the fastest when it is deformed in the direction of the normal vector.

[16] That is, $f(0)=f'(0)=0$. This can be assumed because any surface can be locally represented as a graph over its tangent plane at a point, so a change of coordinates can send the tangent plane to the xy-plane.

To see this more formally, let $\phi : \mathcal{U} \to \mathbb{R}^3$ parameterize S and $f : \mathcal{U} \to \mathbb{R}$ be a function. Then, a deformation of S is parameterized by $\phi_\varepsilon = \phi + \varepsilon f \cdot N$, with ε sufficiently small. Then, letting $g_\varepsilon(u) = [D\phi_\varepsilon]_u^{\mathrm{T}}[D\phi_\varepsilon]_u$ and $g = g_0$,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\operatorname{Area}(\phi_e(\mathcal{U}))\bigg|_{\varepsilon=0} &= \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\int_{\mathcal{U}}\sqrt{\det(g_e(u))}\,\mathrm{d}u\bigg|_{\varepsilon=0}\\ &= \frac{1}{2}\int_{\mathcal{U}}\operatorname{Tr}\left(g^{-1}\,\frac{\mathrm{d}g_\varepsilon(u)}{\mathrm{d}\varepsilon}\bigg|_{\varepsilon=0}\right)\sqrt{\det(g(u))}\,\mathrm{d}u\\ &= -\int_{\mathcal{U}} H(u)f(u)\sqrt{\det(g(u))}\,\mathrm{d}u.\end{aligned}$$

This depends on a nice linear-algebraic fact that the derivative of the determinant of a matrix is its trace.

Moving to the induced metric, notice that the above g has appeared a lot. Maybe that means it's important.

Definition. The object $g = [D\phi_u]^{\mathrm{T}}D\phi_u$ is called the induced metric of a surface S parameterized by $\phi : \mathcal{U} \to \mathbb{R}^3$.

If $E_i = \frac{\partial\phi}{\partial u^i}(0)$ are the tangent vectors of S at $\phi(u)$, then the components of g are $g_{ij} = E_i^{\mathrm{T}}E_j = \langle E_i, E_j\rangle$. Thus, the indiced metric is the restriction of the Euclidean inner product to $T_{\phi(u)}S$, pulled back to $\mathcal{U}$ by ϕ. It's not the inner product on $\mathbb{R}^3$, but it lives in the parameter plane. Thus, a parameterization gives a representation of the induced metric as a matrix.[17]

Is the induced metric independent of parameterization? It's a little more complicated than in the case of scalar quantities. For vector or tensor quantities, the components computed in two different parameterizations may be different. This is because the change in basis affects this, and must be taken into account. Thus, there are transformation formulas for passing from one set of coordinates to another. This notion is called covariance.

The metric tensor is covariant: let $\phi : \mathcal{U} \to \mathbb{R}^3$ and $\psi : \mathcal{V} \to \mathbb{R}^3$ both parameterize S with $\phi(u) = \psi(v) = p \in S$. Let $e_1, \dots, e_n$ be the standard basis vectors in $\mathcal{U}$ and $f_1, \dots, f_n$ be those in $\mathcal{V}$, and let $E_i = \frac{\mathrm{d}\phi}{\mathrm{d}u^i} = D\phi_u \cdot e_i$ and $F_i = \frac{\partial\psi}{\partial v^i} = D\psi_v \cdot f_i$, so that $E_1, \dots, E_n$ and $F_1, \dots, F_n$ are both bases of T_pS. Then,

$$F_i = \frac{\mathrm{d}\psi}{\mathrm{d}v^i} = \sum_v \frac{\partial u^j}{\partial v^i}E_j,$$

so

$$\langle F_k, F_\ell\rangle = \left\langle \sum_i \frac{\partial u^i}{\partial v^k}E_i, \sum_j \frac{\partial u^j}{\partial v^\ell}E_j\right\rangle = \sum_{ij}\frac{\partial u^i}{\partial v^k}\frac{\partial u^j}{\partial v^\ell}g_{ij}.$$

This is what it means to be covariant, and every vector quantity should be checked for covariance at formulation, just as scalar quantities should be checked for independence of parameterization. The metric tensor leads to all sorts of intrinsic geometry, but that's for later.

Turning to geodesics, the basic question one wishes to answer is the shortest path between any two paths on a surface. This is an example of intrinsic geometry, and has some interesting relations to the induced metric. It can be found to satisfy an equation: let $\gamma : I \to S$ be a curve representing the shortest path, and let γ_ε be a variation with endpoints fixed. Then, $\|\dot\gamma_\varepsilon\|$ is an intrinsic quantity, since $\dot\gamma_\varepsilon$ is tangent to the curve. Thus, the length of γ is a minimum of the arc length function, or

$$0 = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\operatorname{Length}(\gamma_\varepsilon)\bigg|_{\varepsilon=0} = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\int_I \|\dot\gamma_\varepsilon(t)\|\,\mathrm{d}t\bigg|_{\varepsilon=0}$$

over all variations. From the homework, this implies that $0 = \langle \mathbf{k}_\gamma, V\rangle$, where V is the variation vector field, or the field that keeps the variations within S, so thst V msut be tangent to S. If $\mathbf{k}_\gamma$ satisfies this, then it is orthogonal to S. Any curve that satisfies that equation or condition is called a geodesic. Note that we don't yet know that these are the shortest paths, just that there is some sort of relation.

The geodesic equation above is a second-order ODE for γ, so there exists a unique local solution for every choice of $p = \gamma(0) \in S$ and $X = \dot\gamma(0) \in T_pS$ (since a 2^{nd}-order ODE needs two initial conditions).

Definition. The choice of (p, X) to a solution at distance 1 is called the geodesic exponential map, denoted $\exp_p : B_\varepsilon(0) \subseteq T_pS \to S$, where $\exp_p(X)$ is one unit of arc-length along the geodesic γ such that $\gamma(0) = p$ and $\dot\gamma(0) = X$. Sometimes, this is also called the geodesic flow.

Then, the geodesic itself is given by $\gamma(t) = \exp_p(tX)$. The exponential map will be useful for the local analysis of geodesics. This is strictly local, because the existence theorem for ODEs don't guarantee that the domain of the exponential map is the entire surface (unless the surface is well-behaved, as in when it is compact). This tool is used to prove useful properties about the geodesics.

[17] Actually, it's a $(2,0)$-tensor, but that will be explained later.

Proposition 9.1. $\exp_p$ *is locally a diffeomorphism near the origin in* T_pS.

Proof. It is easy to see that $[D\exp_p]_0 = \mathrm{id}$, from which the proposition follows. ⊠

Proposition 9.2 (Gauss Lemma)**.** *Let* $v, w \in T_v(T_p(S))$. *Then,* $\left\langle [D\exp_p]_v(v), [D\exp_p]_v(w) \right\rangle = \langle v, w \rangle$.

And then, the most important consequence: geodesics locally minimize length.

Theorem 9.3. *If* γ *is a sufficiently short geodesic and* c *is a curve with the same endpoints as* γ, *then* $\mathrm{Length}(\gamma) \leq \mathrm{Length}(c)$, *with equality iff* $\gamma = c$.

See the supplements for the proofs.

Thus, we have:

- Length-minimizing curves are geodesics.
- Short geodesics are length-minimizing.
- There exist long geodesics which aren't length-minimizing. (For example, on a sphere, geodesics are given by great circles. Two points that are close by have two paths on the circle, one much shorter than the other.)

Thus, the surface S can be turned into a metric space with distance given by $d(p,q) = \inf_\gamma \mathrm{Length}(\gamma)$ for all γ from p to q. Then, it can be shown that d is continuous and satisfies the triangle inequality.

Theorem 9.4 (Hopf-Rinow)**.** *If the exponential map is globally defined, then any two points* p, q *can be connected by a (length-minimizing) geodesic of length* $d(p,q)$.

10. 5/1/13

11. Covariant Differentiation: 5/6/13

"This isn't us [erasing blackboard]... oh wait, it is? It said 468!"

Up to this point, the course has dealt with a lot of extrinsic geometry: the variation of the unit normal vector field, the second fundamental form, and the curvature. However, there is also intrinsic geometry: the induced metric (i.e. the Euclidean metric restricted to the tangent planes) and its parameterization $g_u = [D\phi_u]^{\mathrm{T}}[D\phi_u]$. This can be used to express intrinsic lengths: $\ell = \sqrt{[\dot c(t)]^{\mathrm{T}} \cdot g_c(t) \cdot [\dot c(t)]}\,\mathrm{d}t$. This is the inner product with respect to the metric g.

There is a loose end of sorts in the intrinsic geometry story so far: the geodesic equation for length-minimizing curves is $\mathbf{k}_\gamma(t) \perp T_{\gamma(t)}S$, which is completely extrinsic. It turns out that it will be possible to express the geodesic equation in terms of g alone, so that it is an intrinsic property, which will be a system of 2^{nd}-order differential equations (which incidentally shows the existence of geodesics), but will require the notion of covariant differentiation on S.

Multivariate differentiation in $\mathbb{R}^3$ is associated with a direction. Let $V = [V^1, V^2, V^3]^{\mathrm{T}}$ be a vector in $T_p\mathbb{R}^3$ and let $c : I \to \mathbb{R}^3$ be a curve such that $c(0) = p$ and $\dot c(0) = p$. Then, the derivative of a scalar function f in the direction of V is

$$D_V f = \left.\frac{\mathrm{d}f(c(t))}{\mathrm{d}t}\right|_{t=0} = \sum_{i=1}^{3} V^i \cdot \frac{\partial f}{\partial x^i} = [Df_p]^{\mathrm{T}} \cdot V.$$

Then, the derivative of a vector field is just the above equation applied to the components: let $Y(x) = [Y^1(x), Y^2(x), Y^3(x)]^{\mathrm{T}}$ is given by $D_V Y = [D_V Y^1, D_V Y^2, D_V Y^3]^{\mathrm{T}}$.

On a surface, there's no reason things should be any different for curves: if $f : S \to \mathbb{R}$ and $p \in S$, then it is still possible to define $D_V f = \frac{\mathrm{d}}{\mathrm{d}t} f(c(t))\big|_{t=0}$, where $c : I \to S$ satisfies $c(0) = p$ and $\dot c(0) = V$. However, if $Y : S \to TS$ is a vector field, then the formula given above doesn't necessarily yield a tangent field to S. Basically, the difference of two tangent vectors is not guaranteed to be a tangent vector.

Thus, what sort of alternatives might exist? Ideally, there would be a geometric definition, or failing that, something that might be given by a parameterization that happens to be independent of parameterization. This would require differentiating the coordinate vectors, which makes establishing independence of parameterization difficult. In particular, the basis is different from point to point, which makes life interesting.

We can begin with a geometric definition.

Definition. The covariant derivative of a vector field Y on S in the direction of a $V \in T_pS$ is $\nabla_V Y = [D_V Y]^{\parallel}$.

This is just the component that is parallel to S, so this is useful. Here, $D_V Y$ is the Euclidean derivative given above. Additionally, this is related to the second fundamental form: $D_V Y = [D_V Y]^{\perp} + [D_V Y]^{\parallel} = A(V,Y) \cdot N + \nabla_V Y$. This is because $\langle D_V Y, N \rangle = D_V \langle Y, N \rangle - \langle Y, D_V N \rangle = A(Y,V)$ because the first part goes to zero.

This definition of $\nabla_V Y$ depends only on the first-orders of V and Y along the curve c, which means that it's an extremely local property. Additionally, there are five useful properties, which show that it "looks like" a derivative:

(1) The covariant derivative is C^∞-linear in the V slot: $\nabla_{V_1+fV_2}Y = \nabla_{V_1}Y + f\nabla_{V_2}Y$, where $f : S \to \mathbb{R}$. This property seems reasonable, but underlies a difficult question: the covariant derivative is given by a curve, but what is the curve for $V_1 + fV_2$? It can be done via a parameterization, so it's not actually an issue, but it is worth considering.
(2) The covariant derivative is $\mathbb{R}$-linear in Y: $\nabla_V(Y_1 + a_Y 2) = \nabla_V Y_1 = a\nabla_V Y_2$ for any $a \in \mathbb{R}$.
(3) The product rule: if $f : S \to \mathbb{R}$, then $\nabla_V(fY) = f \cdot \nabla_V Y + (D_V f) \cdot Y$.
(4) There's actually a generalization of the product rule called the metric compatibility property: $D_V \langle Y, Z\rangle = \langle D_V Y, Z\rangle + \langle Y, D_V Z\rangle$. This can be seen because it holds in Euclidean space, following from the product rule. This means that the covariant derivative interacts in the nicest possible way with the inner product.
(5) The covariant derivative is torsion-free: $\nabla_{V_1}\nabla_{V_2} - \nabla_{V_2}\nabla_{V_1})Y = \nabla_{[V_1,V_2]}Y$, where $[V_1, V_2]$ is the Lie bracket (commutator) of two vector fields. This is tangent to S if V_1 and V_2 are. Intuitively, the mixed partials of any C^2 function commute, so $\frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j} - \frac{\partial}{\partial x^j}\frac{\partial}{\partial x^i} = 0$ as an operator on scalar function. This property is the analogous result for vector fields.

Now, it is possible to study covariant derivatives from the parameter domain. Let $\phi : \mathcal{U} \to S$ be a parameterization such that $\phi(0) = p$. A local basis for the tangent planes $T_{\phi(u)}S$ near p is given by the parameterization: $E_i(u) = \frac{\partial\phi}{\partial u^i}$. Then, define the Christoffel symbols as the components Γ^k_{ij} in the following equation:

$$\nabla_{E_i}E_j = \sum_k \Gamma^k_{ij}E_k.$$

Then, if $V = \sum_i a^i E_i(0)$ and $Y_{\phi(u)} = \sum_i b^i(u)E_i(u)$. Using the properties above, the covariant derivative can be computed:

$$\begin{aligned}
\nabla_V Y &= \nabla_{\sum_i a^i E_i}\left(\sum_j b^j E_j\right) \\
&= \sum_{i,j} a^i \nabla_{E_i}(b^j(u)E_j(u)) \\
&= \sum_{i,j} a^i(\nabla_{E_i}(b^j)E_j + a_i b_j \nabla_{E_i}E_j) \\
&= \sum_k \left(\sum_i a^i \frac{\partial b^k}{\partial u^i} + \sum_{i,j} a^i b^j \Gamma^k_{ij}\right) E_k.
\end{aligned}$$

The last step is accomplished by expressing the covariant derivative in the E-basis. This is possible because it is known to be tangent to S.

Lemma 11.1 (Fundamental Lemma of Riemannian Geometry)**.** *The covariant derivative is uniquely defined given an induced metric g that satisfies the five properties above. The relationship between the two is given by*

$$\Gamma_{ijk} = \frac{1}{2}\left(\frac{\partial g_{ik}}{\partial u^j} + \frac{\partial g_{kj}}{\partial u^i} - \frac{\partial g_{ij}}{\partial u^k}\right),$$

where $\Gamma_{ijk} = g(\nabla_{E_i}, E_j, E_k)$.

This is also known as the Levi-Civita connection.

This can be used to express the curvature intrinsically: $\mathbf{k}_\gamma(t) \perp T_{\gamma(t)}S$ is equivalent to $\nabla_{\dot\gamma}\dot\gamma(t) = 0$, or the covariant second derivative is equal to zero. Thus, one obtains the second-order ODE

$$\frac{\mathrm{d}^2\gamma^k}{\mathrm{d}t^2} + \frac{\mathrm{d}\gamma^i}{\mathrm{d}t}\frac{\mathrm{d}\gamma^j}{\mathrm{d}t} = 0,$$

which holds true when γ is a basis.

Though the gradient is typically defined in terms of level sets, it can be thought of more geometrically thanks to directional derivatives:

Definition. Let $c : I \to S$ be a curve with $c(0) = p$ and $\dot c(0) = V$. Then, the gradient $\nabla f(p)$ is defined by $\left.\frac{\mathrm{d}f(c(t))}{\mathrm{d}t}\right|_{t=0} = \langle\nabla f(p), V\rangle$.

This makes it very clear which direction V must be in to maximize ∇f. In Euclidean space, $\nabla f(p) = [Df_p]^{\mathrm{T}}$, but in the parameter domain this doesn't always work (orthogonality is more complicated), so $\nabla f = g^{-1} \cdot Df$, which is because g comes from the inner product. This might not be orthogonal in $\mathcal{U}$, but it will be in S.

This principle also extends to the other differential operators: in essence, all three carry over to surfaces, but because the notion of orthogonality changes, then g is involved.

- The divergence is given by $\nabla \cdot X = \sum_j \langle \nabla_{E_i} X, E_i \rangle$, where E_i is an orthonormal basis, which happens to be independent of choice of basis and in the parameter domain,
$$\nabla \cdot X = \sum_i \left[\frac{\partial x^i}{\partial u^i} + \sum_j \Gamma^i_{ij} X_j \right].$$
- The Laplacian is given by $\Delta f = \nabla \cdot (\nabla f)$, which can be expressed as
$$\Delta f = \sum_{i,j} g^{ij} \left[\frac{\partial^2 f}{\partial u^i \partial u^j} + \Gamma^k_{ij} \frac{\partial f}{\partial u^k} \right].$$

These will be useful in the next lecture.

12. Discrete Laplacians: 5/8/13

The goal of today is to derive the Laplacian, which is an extremely important operator. However, this is defined differently in the math and computer science circles and thus one should be wary of sign errors in today's lecture.

The Laplacian is a linear functional, which is in some sense a generalization of a matrix. Specifically, $\Delta : C^\infty(M) \to C^\infty(M)$, so it acts on a space of functions. On some domain $\Omega \subseteq \mathbb{R}^n$, the Laplacian operator takes on the form $\Delta = -\sum_i \frac{\partial^2}{\partial x_i^2}$. The Laplacian shows up in a lot of physics: the heat equation, one of the most important PDEs, is just $\frac{\partial u}{\partial t} = -\Delta u$. In some sense, this represents how a heat distribution tends to equalize.

Another useful thing to do is take the eigenfunctions $\Delta f = \lambda f$ (eigenvectors in some sense, but on the function space). An object might have certain vibration modes or resonant frequencies, and these are given by the eigenvalues.

The easiest way to define the Laplacian is $\Delta f = \operatorname{div} \operatorname{grad} f$. This hides some other definitions, and if g is the induced metric, this becomes
$$\Delta f = \frac{1}{\sqrt{|g|}} \partial_i \left(\sqrt{|g|} g^{ij} \partial_j f \right). \tag{2}$$
Unsurprisingly, this is difficult to discretize. The gradient isn't so bad, but the divergence will be harder. Thus, an alternate definition will be used. This requires a slight generalization of integration by parts: if $\Omega \subseteq \mathbb{R}^n$ is a domain, then
$$\int_\Omega f \Delta g \, \mathrm{d}A = \text{ some boundary terms } - \int_\Omega \nabla f \cdot \nabla g \, \mathrm{d}A.$$
Stokes' theorem can also be used to justify the result.

Let $f : M \to \mathbb{R}$. Then, one can generate an operator called the L^2 dual to f, $\mathcal{L}_f : L^2(M) \to \mathbb{R}$, where $\mathcal{L}_f[g] = \int_M fg \, \mathrm{d}A$. Here, g is called a test function, and $L^2(M)$ is the space of square-integrable functions $f : M \to \mathbb{R}$ (i.e. those such that $\int_M f^2 \, \mathrm{d}A$ is a real number). The requirement that f is square-integrable comes from a definition of the inner product on a function space, $\langle f, g \rangle = \int fg$, so $\langle f, f \rangle = \|f^2\| = \int f^2$. The set of test functions is $\{g \in C^\infty(M) \mid g_{\partial M} = 0\}$[18] In some sense, this is the set of functions that one might wish to take the Laplacian of. Finally, by restricting g to a very small domain, it is possible to recover f at a given point from its dual.

This can be used to skip the notion of divergence entirely: suppose f is infinitely differentiable. Then
$$\mathcal{L}_{\Delta f}[fg] = \int_M g \Delta f \, \mathrm{d}a = -\int_M \nabla g \cdot \nabla f \, \mathrm{d}A,$$
since $g = 0$ on ∂M, so the boundary terms in (2) vanish. This is an example of Galerkin's approach to PDEs. The key insight for this weak Laplacian is that only one derivative of f and g is necessary, sothey only have to be C^1.

One can obtain a basis for the space of functions given by the hat functions $h_i(x)$ as in Figure 2. Then, $f(x) = \sum_i a_i h_i(x)$, given by some $a \in \mathbb{R}^{|V|}$, where V is a set of vertices. Thus, if one represents two fucntions f and g as linear combinations of hats, the expression $\nabla f \cdot \nabla g$ will be constant on each face.

In order to actually evaluate this, let h be a hat function: $h(v_1) = 1$ and $h(x) = 0$ if $x \neq v_1$. Then, let v_2 and v_3 be the other vertices on a triangle with v_1. Since the gradient is perpendicular to the surface, then $\nabla h \cdot n = 0$, and since the gradient is linear along the edges, $\nabla h \cdot (v_1 - v_3) = 1$, and similarly $\nabla h \cdot (v_1 - v_2) = 1$. Thus, $\nabla h \cdot (v_2 - v_3) = 0$, so it must lie on the triangle, pointing from the v_2v_3-edge towards v_1.

Let θ_3 be the angle out of v_3 and ℓ_3 be the length of the v_1v_3-edge. Then, $1 = \nabla h \cdot (v_1 - v_3) = \|\nabla f\| \ell_3 \cos(\pi/2 - \theta_3) = \|nabla f\| \ell_3 \sin\theta_3$, so $\|\nabla f\| = 1/(\ell_3 \sin\theta_3) = 1/h$, where h is the height of the triangle. Alternatively, let e_{23} be the e_2e_3-edge and A be the area, so that $\nabla f = e_{23}^\perp / 2A$ (e_{23} is rotated 90°).

[18] Sometimes, $\partial M = \emptyset$, so the last condition is vacuously true.

Now, it's necessary to compute $\nabla f \cdot \nabla g$. If f and g are hat functions, then there are two cases:

- If f and g are hat functions on the same vertex, then on the triangle T,
$$\int_T \langle \nabla F, \nabla F \rangle = A\|\nabla F\|^2 = \frac{A}{h^2} = \frac{b}{2h} = \frac{1}{2}(\cot\alpha + \cot\beta),$$
as shown in a previous lecture. The variable names are what you might expect, though α and β are the angles at vertices v_2 and v_3,
- If they are on adjacent but different vertices, call the functions f_α and f_β. Then,
$$\begin{aligned}\int_T \langle \nabla f_\alpha, f_\beta\rangle &= A\langle \nabla f_\alpha, f_\beta\rangle \\ &= \frac{1}{4A}\left\langle e_{31}^\perp, e_{12}^\perp\right\rangle = \frac{-\ell_1\ell_2\cos\theta}{4A} \\ &= \frac{h^2\cos\theta}{4A\sin\alpha\sin\beta} = \frac{-\cos\theta}{2b\sin\alpha\sin\beta} \\ &= -\frac{\cos\theta}{2\sin(\alpha+\beta)} = -\frac{\cot\theta}{2},\end{aligned}$$
where θ is the angle of vertex v_1, and ℓ_i is the length of the edge between v_1 and v_i.
- Otherwise, the inner product is zero.

Then, one can get a more general formula by summing around a vertex. This should look like the surface area formula defined earlier, which is not a coincidence. Thus, one obtains a formula called the cotangent Laplacian:

$$L_{ij} = \begin{cases} \frac{1}{2}\sum_{i\sim j}(\cot\alpha_j + \cot\beta_j), & i = j \\ -\frac{1}{2}\cot(\alpha_i + \beta_j) & i \sim j \\ 0, & \text{otherwise.}\end{cases}$$

The Poisson equation is $\Delta f = g$, where the problem is to find f given g. This is called a strong equation, since it's in terms of the derivative of f, but it can be put into the weak (integrated) form. A function is a weak solution to a Poisson equation if $\int_M \phi\Delta f\,\mathrm{d}A = \int_m \phi g\,\mathrm{d}A$ for all test functions ϕ. All solutions are weak solutions, at least.

In the finite case, this becomes $\int_M h_i\Delta f\,\mathrm{d}A = \int_M h_i g\,\mathrm{d}A$ for all hat functions h_i. However, the Laplacian can be expanded:

$$\int_M h_i\Delta f\,\mathrm{d}A = -\int_M \nabla h_i\cdot\nabla f\,\mathrm{d}A = -\int_M \nabla h_i\cdot\nabla\left(\sum_j a_j h_j\right)\mathrm{d}A = \sum_j L_{ij}a_j.$$

Thus, this becomes multiplying by a Laplacian matrix, which is relatively nice. However, there is a bit of a problem: the right-hand side of this weak solution requires multiplying two piecewise-linear functions, which is quadratic. This is harder to work with, but there are several approaches:

(1) The consistent approach is to just do the integral.
(2) Alternatively, one can make some more approximations.
(3) A third approach is to slightly redefine g.

Definition. The mass matrix of a triangle mesh is

$$A_{ij} = \int_M h_i h_j\,\mathrm{d}A.$$

The diagonal element $a_{ii} = \|h_i\|$, and the off-diagonal elements a_{ij} represent the overlap between h_i and h_j. In order to implment this in any reasonable manner, it is necessary to represent this as a sparse matrix.

Using the consistent approach (i.e. choice 1), one can take the total area in the fan and divide by 6 (if $i = j$) or 12 (if $i \neq j$). Then, the rows of the mass matrix sum to the one ring area divided by 3, which means that in some sense this partitions the surface area (because each face touches 3 vertices).

The mass matrix can also be used for integration:

$$\int_M f = \int_M \sum_j a_j h_j = \int_M \sum_j a_j h_j \sum_i hi = \sum_{ij} A_{ij}a_j = \mathbf{1}^{\mathrm{T}}A\mathbf{a}.$$

Since $\mathbf{1}^{\mathrm{T}}A\mathbf{1}$ is the surface area, there is a clear relation. However, this matrix isn't diagonal, which is unpleasant. It can be approximated with a lumped mass matrix, which takes all of the off-diagonal values and put them in the nearest diagonal, so that $\widetilde{a}_{ii}$ is the area of the i^{th} cell in the dual mesh. For smooth functions, this is a good approximation, since there isn't too much variation between neighboring vertices. This leads to an argument for convergence: as the triangle mesh is refined, this becomes exact in the limit. The way to associate area to a vertex is

to associate one-third of each triangle (divided by the barycenter) to each vertex. This given a cell which isn't always convex, which is admittedly pretty suspicious. Other approximations exist, and have their own drawbacks.

Solving the Poisson equation can be done with a computer without too much difficulty, thanks to the cotangent Laplacian. Second-order differential operators correspond to positive-definite quadratic forms, so by something in PDE called elliptic regularity, these solutions are guaranteed to be smooth. There are lots of other useful things that can be done with this operator. One can also use the heat equation to smooth out a triangle mesh. This is one line in Matlab once the cotangent Laplacian is implemented, and helps make noisy scans smoother.

13. Tensors and Exterior Calculus: 5/13/13

"Some previous lecturer — I hope it wasn't me — walked off with all of the chalk."

Definition. If $\mathcal{V}$ is an n-dimensional vector space, then an inner product on $\mathcal{V}$ is a bilinear, symmetric, positive-definite function $\langle \cdot, \cdot \rangle : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$.

By bilinear we mean that $\langle v + aw, x \rangle = \langle v, x \rangle + a \langle w, x \rangle$ and $\langle v, w + ax \rangle = \langle v, w \rangle + a \langle v, x \rangle$ (i.e. linear in both arguments). Symmetry is just that $\langle v, w \rangle = \langle w, v \rangle$, and positive-definite means that $\langle v, v \rangle \geq 0$ and is zero only when v is.

This should be thought of as defined on an abstract vector space, but induces lots of familiar linear-algebraic terms:

(1) The norm of a vector is $\|v\| = \sqrt{\langle v, v \rangle}$.
(2) Two vectors $v, w \in \mathcal{V}$ are orthogonal if $\langle v, w \rangle = 0$.
(3) The notion of orthogonal complement follows: if $\mathcal{S}$ is a subspace of $\mathcal{V}$, then $\mathcal{V} = \mathcal{S} \oplus \mathcal{S}^{\perp}$.

Definition. If $\mathcal{V}$ is a vector space, then the dual space is $\mathcal{V}^* = \{\xi : \mathcal{V} \to \mathbb{R}, \text{ where } \xi \text{ is linear}\}$. The elements of $\mathcal{V}^*$ are called linear functionals.

Proposition 13.1. *$\mathcal{V}^*$ is an n-dimensional vector space.*

Proof. It's not hard to see that this is a vector space: if ξ and ζ are linear functionals, it's clear that $\xi + \zeta$ and $a\xi$ are both linear functionals.

To show that it has dimension n (where $n = \dim(V)$), suppose $\{E_i\}$ is a basis for V. Then, let $\omega^i(E_s) = 1$ if $i = s$ and be zero otherwise. This is clearly linearly independent, and to show that it spans $\mathcal{V}^*$ show that any $\xi \in \mathcal{V}^*$ can be written as $\xi = \sum_{i=1}^{n} \xi(E_i)\omega^i$. ⊠

If $\mathcal{V}$ is an inner product space, there are three additional constructions:

- The flat operator: if $v \in \mathcal{V}$, then $v^\flat \in \mathcal{V}^*$ is given by $v^\flat(W) = \langle v, w \rangle$ for all $w \in \mathcal{V}$. In some sense, $\flat : v \mapsto \langle v, \cdot \rangle$. $v^\flat$ is read 'v-flat."
- Somewhat unsurprisingly, the sharp operator is the inverse operation: if $\xi \in \mathcal{V}^*$, then there exists a $\xi^\sharp \in \mathcal{V}$ such that $\xi(w) = \langle \xi^\sharp, w \rangle$ for any $w \in \mathcal{V}$. Notice that $(\xi^\sharp)^\flat = \xi$ and $(v^\flat)^\sharp = v$.[19]
- $\mathcal{V}^*$ carries an induced inner product $\langle \xi, \zeta \rangle_{\mathcal{V}^*} = \langle \xi^\sharp, \zeta^\sharp \rangle_{\mathcal{V}}$.

The existence of the flat and sharp operators is a special case of the Riesz representation theorem, which guarantees an existence and uniqueness result like that of the sharp operator in all sorts of infinite-dimensional spaces. This has applications in PDEs.

The fact that no basis is necessary here is nice; sometimes (e.g. infinite-dimensional spaces) it's unpleasant to deal with, and definitions that are independent of basis are also nicer. However, if there is a basis $\{E_i\}$ for $\mathcal{V}$, then let $g_{ij} = \langle E_i, E_j \rangle$ and let g^{ij} be the components of the inverse matrix to $[g_{ij}]$.[20] Then, the following can be computed:

- The dual basis of $\mathcal{V}^*$ is $\omega^i = \sum_j g^{ij} E_j^\flat$, which can be seen because

$$\omega^i(E_s) = \sum_j g^{ij} E_j^\flat(E_s) = \sum_j g^{ij} \langle E_j, E_s \rangle = \sum_j g^{ij} g_{js} = \delta_j^i,$$

  where δ_j^i is 1 if $i = j$ and is zero otherwise.
- If $v = \sum_i v^i E_i$, then $v^\flat = \sum_i v_i \omega^i$, where $v_i = \sum_j g_{ij} v^j$.
- Similarly, if $\xi = \sum_i f_i \omega^i$, then $f^\sharp = \sum_i f^i E_i$, where $f^i = \sum_j g^{ij} f_j$.
- If $\xi = \sum_i a_i \omega^i$ and $\zeta = \sum_i b_i \omega^i$, then $\langle \xi, \zeta \rangle = \sum_{i,j} g^{ij} a_i b_j$.

[19] Amusingly enough, these are *natural* transformations, since they are defined independent of basis. However, showing all of this does require a proof.

[20] Aside on indices: convention dictates where indices are placed on things: if v is a vector, then the vector indices are upstairs and vector basis indices are downstairs: $v = \sum v^i E_i$. However, if ξ is a linear functional, the dual vector indices are written *down*stairs and the dual basis indices are written *up*stairs, or $\xi = \sum \xi_i \omega^i$. The example g_{ij} given above is a "dual" object that acts on two vectors, and its components go downstairs, so one gets g_{ij}.

Because g_{ij} and g^{ij} move indices up and down the stairs, they are sometims known as raising and lowering opertations (especially in relativity), which just means that they perform the flat and sharp operations. If E_i is orthonormal most of these simplify nicely.

Definition. If $\mathcal{V}$ is a vector space of dimension n, a k-covariant and ℓ-contravariant tensor is a multilinear function $f : \mathcal{V}^k \to \mathcal{V}^\ell$.

In some sense, this creates a multi-vector. The space of k-covariant and ℓ-contravariant tensors is a vector space of dimension $n^{k+\ell}$, written $(\mathcal{V}^*)^k \otimes \mathcal{V}^\ell$, and it has a basis in terms of the E_i and ω^i. Additionally, if $\mathcal{V}$ is an inner product space, the tensor space is as well, and sharp and flat operators exist, but there is one operator for each index. Additionally, there are contraction operations, where a $\mathcal{V}$-term and a $\mathcal{V}^*$-term are both killed: a (k, ℓ)-tensor is sent to a $(k-1, \ell-1)$-tensor. Note that the metric is an example of a $(k, 0)$-tensor.

A symmetric $(2, 0)$-tensor is an $A \in \mathcal{V}^* \otimes \mathcal{V}^*$ such that $A(v, w) = A(w, v)$ for all $v, w \in \mathcal{V}$. The second fundamental form and the shape operator are both symmetric $(2, 0)$-tensors.

- If there is a basis $\{\omega^i\}$ for $\mathcal{V}^*$, then there is a basis $\{\omega^i \otimes \omega^j\}$ of $\mathcal{V}^* \otimes \mathcal{V}^*$, where $\omega^i \otimes \omega^j(v, w) = \omega^i(v)\omega^j(w)$. In this basis, $A(v, w) = \sum_{i,j} A_{ij}\omega^i \otimes \omega^j$, and $A_{ij} = A_{ji}$. Then,

$$A(v, w) = \sum A_{ij}\omega^i(v)\omega^j(w) = \sum A_{ij}\omega^i(E_s)w^s\omega^j(E_t)w^t = \sum A_{ij}\delta^i_s v^s \delta^j_t w^t = \sum_{s,t} v^s w^t A_{st}.$$

  This can also be written as $[v]^{\mathrm{T}}[A][w]$.
- It is possible to define an associated self-adjoint $(1, 1)$-tensor $S \in \mathcal{V}^* \otimes \mathcal{V}$ given by $A(v, w) = \langle S(v), w\rangle$. This is how the shape operator was obtained from the second fundamental form. If a basis is given, then $S = \sum_{i,j} S^j_i \omega^i \otimes E_j$, where $S^j_i = \sum_k g^{kj} A_{ik}$.
- If $v = \sum_i v^i E_i$ and $w = \sum_i w^i E_i$, then $\langle v, w\rangle = v^{\mathrm{T}} g w$, $A(v, w) = v^{\mathrm{T}} A w$, and $S = g^{-1}A$.
- The contraction of A is $\mathrm{Tr}(S) = \sum_{i,j} g^{ij} A_{ij}$.

An alternating $(k, 0)$-tensor, also called a k-form, is a $\sigma \in \mathcal{V}^* \otimes \cdots \otimes \mathcal{V}^*$ such that for any $v, w \in \mathcal{V}$ and pairs of slots in σ, $\sigma(\dots, v, \dots, w, \dots) = -\sigma(\dots, w, \dots, v, \dots)$ (i.e. switching them inverts the sign).

If $\dim \mathcal{V} = 2$, then $\mathrm{Alt}^0(V) = \mathbb{R}$ (zero-tensors are just scalars), $\mathrm{Alt}^1(V) = \mathcal{V}^*$, and $\mathrm{Alt}^2(\mathcal{V}) \cong \mathbb{R}$, and any higher dimensions are zero: if $v, w, x \in \mathbb{R}^2$, then $x = a + bw$, so $\sigma(v, w, x) = a\sigma(v, w, v) + b\sigma(v, w, w) = 0$, because v can be switched with itself in the first term, leaving the expression unchanged, but also inverting its sign, so it must be zero. The second term is similar. In general, $n+1$-tensors vanish over a vector space of dimension n.

A basis for $\mathrm{Alt}^2(\mathcal{V})$ can be given by $\omega^1 \wedge \omega^2$, defined as $\omega^1 \wedge \omega^2(v, w) = \det([v\ w])$.

Definition.

- If $\mathcal{V}$ is an inner product space, the area form $\mathrm{d}A \in \mathrm{Alt}^2(\mathcal{V})$ is given by $\mathrm{d}A(v, w)$ is equal to the signed area of the parallelogram $v \wedge w$.
- The Hodge-star operator is $\omega \wedge \star\tau = \langle \omega, \tau\rangle\ \mathrm{d}A$.

Here, $\star\,\mathrm{d}A = 1$ and $\star 1 = \mathrm{d}A$, and this can be thought of sort of as rotation by $90°$ in some contexts. The Hodge star is actually relatively nice in the discrete case, and more intuition will be given in a future lecture.

Returning to differential geometry, let $\mathcal{V}_p = T_p S$.

Definition. The bundle of (k, ℓ)-tensors over S attaches the vector space $\mathcal{V}_p^{(k,\ell)} = \mathcal{V}_p^* \otimes \cdots \otimes \mathcal{V}_p^* \otimes \mathcal{V}_p \otimes \cdots \otimes \mathcal{V}_p$. A section of this bundle is an assignment $p \mapsto \sigma_p \in \mathcal{V}_p^{(k,\ell)}$.

For example, if $k = \ell = 0$, then the sections are functions on S. A section on a $(0, 1)$-bundle is a vector field on S. A section of a $(1, 0)$-bundle is a choice of a linear functional at each point, called a one-form. A symmetric $(2, 0)$-section includes the metric and the second fundamental form. An antisymmetric $(2, 0)$-form is called a two-form, and includes the area form. The tangent and cotangent bundles are examples of these tensor bundles.

The covariant derivative extends naturally to tensor bundles, which is an example of a more general theme (e.g. you also get flats and sharps). Though this is independent of basis, choose a basis for calculation. The covariant derivative takes $k+1$ vectors and returns ℓ vectors, where the extra input specifies the direction the derivative is taken in. The formula for a tensor

$$\sigma = \sum_{i,j,k,l} \sigma^{kl}_{ij} \omega^i \otimes \omega^j \otimes E_k \otimes E_l.$$

Then, the covariant derivative is

$$\nabla\sigma = \sum_{i,j,k,l,s} \nabla_s \sigma^{kl}_{ij} [\omega^i \otimes \omega^j \otimes E_k \otimes E_{il}] \otimes \omega_s,$$

where

$$\nabla_s \sigma_{ij}^{kl} = \frac{\partial \sigma_{ij}^{kl}}{\partial x^s} - \Gamma_{is}^t \sigma_{tj}^{kl} - \Gamma_{js}^t \sigma_{it}^{kl} + \Gamma_{ts}^i \sigma_{ij}^{tl} + \Gamma_{ts}^l \sigma_{ij}^{kt}.$$

Definition. The exterior derivative $\mathrm{d} : \mathrm{Alt}^k(S) \to \mathrm{Alt}^{k+1}(S)$ is given in some basis by

- $\mathrm{d}f = \sum_i \frac{\partial f}{\partial x^i} \omega^i$ for $f \in \mathrm{Alt}^0(S)$. Thus, $(\mathrm{d}f)^\sharp = \nabla f$.
- If $\omega = \sum_i a_i \omega^i \in \mathrm{Alt}^1(S)$, then $\mathrm{d}\omega = (\frac{\partial a^1}{\partial x^2} - \frac{\partial a^2}{\partial x^1}) \omega^1 \wedge \omega^2$.
- If $\omega \in \mathrm{Alt}^2(S)$, then $\mathrm{d}\omega = 0$.

This means that $\mathrm{d}(\mathrm{d}\omega) = 0$ for any ω, and that the exterior derivative is in some sense the metric-independent part (or antisymmetrization) of the covariant derivative. This is useful for a vast and pretty generalization of the Fundamental Theorem of Calculus:

Theorem 13.2 (Stokes)**.** *If c is a $k+1$-dimensional submanifold of S with k-dimensional boundary ∂c. If ω is a k-form on S, then*

$$\int_c \mathrm{d}\omega = \int_{\partial c} \omega.$$

This implies lots of results in vector calculus, such as the divergence theorem.

Definition. The co-differential δ is the L^2-adjoint of d, and is thus an operator $\delta : \mathrm{Alt}^{k+1}(S) \to \mathrm{Alt}^k(S)$ such that $\int_S \langle \mathrm{d}\omega, \tau \rangle \, \mathrm{d}A = \int_S \langle \omega, \delta\tau \rangle \, \mathrm{d}A$. $\delta = - \star \, \mathrm{d} \star$.

This leads to some familiar concepts: if X is a vector field, then $\delta X^\flat = \mathrm{div}(X)$ and $\mathrm{d}X^\flat = \mathrm{curl}(X) \, \mathrm{d}A$.

Theorem 13.3 (Hodge)**.** $\mathrm{Alt}^1(S) = \mathrm{d}\,\mathrm{Alt}^0(S) \oplus \delta\,\mathrm{Alt}^2(S) \oplus \mathcal{H}^1$*, where $\mathcal{H}^1$ is the set of harmonic one-forms h such that $\mathrm{d}h = \delta h = 0$.*

Sometimes, $(d\delta + \delta d)$ is called the Hodge Laplacian.

Corollary 13.4. *Every vector field X on S can be decomposed into a gradient of some function, a divergence-free part, and a harmonic part: $X = \nabla\phi + \mathrm{curl}(\nabla\psi) + h^\sharp$ for some functions ϕ and ψ and an $h \in \mathcal{H}^1$.*

This leads to another deep mathematical result:

Theorem 13.5. $\dim(\mathcal{H}^1) = 2g(S)$*, where $g(S)$ is the genus of S.*

14. Discrete Exterior Calculus: 5/15/13

The div, grad, and curl operators are related to several theorems from vector calculus:

Theorem 14.1 (Divergence)**.** *Let $\Omega \subseteq \mathbb{R}^2$ be a domain and $\mathbf{v}$ be a vector field. Then,*

$$\int_\Omega \mathrm{div}\, \mathbf{v} \, \mathrm{d}A = \int_{\partial\Omega} \mathbf{v} \cdot \mathbf{n} \, \mathrm{d}l.$$

This says that the total amount that leaves the region can be found by adding up all of the tendency to spread. Looking at circulation, which is relatd to the tangential field, we also have:

Theorem 14.2 (Green)**.** *With Ω and $\mathbf{v}$ as before, let $\mathbf{t}$ be the tangential vector field on $\partial\Omega$. Then,*

$$\int_\Omega \mathrm{curl}\, \mathbf{v} \, \mathrm{d}A = \int_{\partial\Omega} \mathbf{v} \cdot \mathbf{t} \, \mathrm{d}l.$$

This extends to higher dimensions, which is not too hard. But exterior calculus is all about generalizing this to surfaces and manifolds. Since the goal is to be intrinsic, all vector fields will be tangent fields (since an intrinsically defined vector field must be tangent).

Definition. A differential form is an alternating, k-linear functional that at every point p associates a real number with k vectors in the tangent space at p.

For example, if ω is a 2-form, then $\omega_p(v_1, v_2) \to \mathbb{R}$ implies that $v_1, v_2 \in T_pM$, and $\omega_p(v_1, v_2) = -\omega_p(v_2, v_1)$. Recall that a k-form must be zero in an n-dimensional space if $k > n$.

A function $f : \Sigma \to \mathbb{R}$ is a 0-form on Σ, which is a silly example, but allows much of the calculus defined on functions to extend to exterior calculus. A differential one-form is associated with a vector field: if $\mathbf{v} : \Sigma \to T\Sigma$, then $\mathbf{v}^\flat(\mathbf{x}) = \mathbf{v} \cdot \mathbf{x}$ is a differential 1-form, and for each differential 1-form ω, $\omega^\sharp$ is a vector field.[21]

[21] These are called the musical isomorphisms, so named because they raise and lower indices: a sharp raises a note, and $\sharp$ takes a lower index and raises it, and similarly flat lowers notes and $\flat$ gives a lower index from an upper index. These indices happen because there are inner products on the surface and the parameter plane, and these allow one to sort out the difference without having to carrying around the metric everywhere.

There are several other operator: the sharp and flat operators more generally move between covariant and contravariant tensons, and there is an exterior derivative $\mathrm{d}\omega$ and a Hodge star $\star\omega$ which sends k-forms to $(n-k)$-forms and can be thought of as sending a plane to its normal. The wedge product $\omega_1 \wedge \omega_2$, which can be thought of as the cross product, which sends a k-form and an ℓ-form to a $(k+\ell)$-form. Thankfully, the discrete versions of operators are generally easier to understand.

Additionally, a k-form can be integrated on a k-dimensional surface. Intuitively, this measures how much of the differential form is parallel to the surface: if ω is a 1-form and $\gamma : [0,1] \to \Sigma$ is a curve, then $\int_\gamma \omega = \int_0^1 \omega(\gamma')\,\mathrm{d}t = \int_0^1 (\omega^\sharp \cdot \gamma')\,\mathrm{d}t$. This relates to Stokes' theorem, presented in the previous lecture as Theorem 13.2. Notice that this theorem integrates a k-form ω on a k-dimensional object $\partial\Omega$ and a $(k+1)$-form $\mathrm{d}\omega$ on a $(k+1)$-dimensional object Ω. This theorem implies all of the similar-looking results in vector calculus, including Theorems 14.1 and 14.2 above.

This exterior calculus can be discretized in a mostly mathematically rigorous manner, but a couple of approximations could be needed.

Recall the definition of an oriented triangle mesh, which is a topological object that associates a direction to each face (clockwise or counterclockwise) and induced orientations of edges, such that the faces touching each edge have opposite orientations at each edge. Then, the dual to a complex is to replace a k-dimensional object and replace it with an $(n-k)$-dimensional object: replace each triangle with a vertex at its center, and connect them based on adjacency (i.e. if two triangles share an edge, their vertices share an edge). Then, vertices turn into faces, though they aren't always triangles. The edges are rotated by about $90°$ in this dualizing.

Differential forms were made to be integrated, so the trick in the discrete case is to store the integrals of forms. A discrete 0-form is a function, and should be integrated on points, which just returns the function again. Thus, this is a function on vertices, and can be thought of as a vector on $|V|$ components. A discrete 1-form is defined on the edges, so rather than store the form ω, store $\int_e \omega$ for all edges e. Thus, one obtains a vector on $|E|$ components. Notice that the orientation of the edge is important for this calculation. Similarly, a 2-form is defined on the triangular faces, the integrals on each face is stored.

Using Stokes' theorem, one can implement a d operator: if ω is a one-form, then we can calculate $\int_e \mathrm{d}\omega = \int_{\partial e} \omega = \omega_2 - \omega_1$, where ω_1 and ω_2 are the edges of the vertex, with e oriented towards ω_2. Thus, the d operator can be defined only topologically, with the connection to calculus coming later. Notice that no approximation has been made yet. Thus, $\mathrm{d} \in \mathbb{R}^{|E|\times|V|}$ can be thought of as a matrix, and computationally it makes much more sense to store it as a sparse matrix (since most of the entries are 0). However, for ging from 3-forms to 2-forms, there will be a different matrix $\mathrm{d} \in \mathbb{R}^{|F|\times|E|}$, but the concept is the same: suppose ω is a 2-form on a face t, and let $\sigma(e)$ be 1 if the edge has the same orientation as the face, and -1 if not. Then,

$$\int_t \mathrm{d}\omega = \int_{\partial t} \omega = \sigma(\omega_1)\omega_1 + \sigma(\omega_2)\omega_2 + \sigma(\omega_3)\omega_3).$$

Additionally, $\mathrm{d}^2 = 0$, which still holds here, because no approximations have been made yet. However, the operators are different here, so take care.

The Hodge operator can be thought of as moving not just between forms, but from a mesh to its dual mesh. Thus, $\star^2$ is the identity. Then, on the mesh, 2-forms go to 0-forms, and vice versa, and edges (1-forms) go to 1-forms. Here, faces are sent to vertices, but an approximation happens, since the exact location must be guessed.

If one sends a 2-form to a 0-form, the only thing that must be normalized is the area. Thus, the Hodge star is a diagonal matrix with entries $\star_{ii} = 1/A_i$, where A_i is the area of the face i. Then, to change edges into edge, a common choice for the edge $e^\star$ given e is the edge between the circumcenters (so that $e \perp e^\star$) of the two triangles incident on the edge. The ratio of the lengths is

$$\frac{\|e^\star_{ij}\|}{\|e_{ij}\|} = \frac{1}{2}(\cot\alpha_j + \cot\beta_j),$$

where α_j and β_j are the angles opposite the edge e in the triangles it borders. There are various suspicious things going on here, since the adjacent triangles may not be flat against each other, and a sufficiently poorly behaved triangle doesn't contain its circumcenter, making the contangent weight a negative number... that is supposed to represent a ratio of areas. But it converges, so it's okay.

Finally, 0-forms are sent to 2-forms. This is the inverse of the previous operation: $\star_{ii} = A_i$, the area of the i^{th} cell.

Recall that the mass matrix associated areas to the triangle mesh, and since it wasn't diagonal, an approximation called the barycentric lumped mass matrix doesn't work teribly nicely with the dual, since the barycenter and circumcenter aren't necessarily the same. However, there are lots of other options: the Voronoi cell uses circumcenters rather than barycenters, which guarantees convexity, but some points may lie outside of their associated triangle. This can be elegantly defined as the set of points closest to the center vertex compared to any other, which leads to nice convergence properties.

Note that the boundary is a slight headache. Often, the Dirichlet boundary condition means that the boundary is taken to be zero, but in general it is necessary to determine which boundary condition is most appropriate to the application. Often, the mesh is closed, which avoids the whole problem.

Often, a solution called the mixed Voronoi cell is used, which uses either the circumcenter or the midpoint of the circumcenter and the barycenters depending on what's preferable.

Then, one obtains a giant commutative diagram with $\star$, d, and δ called the de Rham complex. Here, $\delta = - \star \, \mathrm{d} \star$, sending k-forms to $(k-1)$-forms. The Hodge Laplacian $\Delta = \mathrm{d} \star \mathrm{d} \star + \star \, \mathrm{d} \star \mathrm{d}$, has lots of nice properties, and most impressively, reduces to the cotangent Laplacian on 0-forms. Thus, the same result happens whether one starts with first-order finite elements or discrete exterior calculus.

One simple application of this is the Helmholtz-Hodge decomposition of vector fields (or 1-forms) into the sum of a divergence-free field, a curl-free piece, and a harmonic piece. This requires decomposing the associated 1-form ω as $\omega = \delta\beta + \mathrm{d}\alpha + \gamma$ where $\mathrm{d}\gamma = \delta\gamma = 0$. Then, $\delta\omega = \delta \, \mathrm{d}\alpha$ and $\mathrm{d}\omega = \mathrm{d}\delta\beta$, and then γ can be found by subtraction. There's an incredibly simple topological algorithm for this, too.

One application of this is to model hurricanes (which have lots of circulation). Then, the place where the divergence vanishes is the eye. Another option is fluid simulation. Fluids are incompressible, so there is no divergence, so the Helmholtz-Hodge decomposition is useful. It can be used to create vector fields, and to measure turbulence in acoustics for computational physics (where it might be framed as a decomposition into solenoidal, or magnetic, and radial, or electric, parts). There are similar applications in optics. There are even applications to machine learning in which an algorithm makes guesses at an entire vector field by guessing the divergene and curl of the vector field. Since the kernels of these fields are useful, this is more applicable than one might think.

15. Isometries and Rigidity: 5/20/13

Recall that the geodesic exponential map of a surface S at a point $p \in S$ is the mapping $\exp_p : T_pS \to S$ such that $\exp_p(V) = \gamma(1)$ where γ is the unique gedoesic through p in the direction V. This depended on the assertion that such a geodesic already exists, but this was found by setting up a system of ODEs for the geodesics, which therefore have a solution. Since the solution may be local, the exponential map may not be well-defined if $\|V\|$ is too large.

By Proposition 9.1, the differential of the exponential is the identity, so it's invertible and therefore there exist open sets $0 \in \mathcal{U} \subseteq T_pS$ and $p \in \mathcal{V} \subseteq S$ such that $\exp_p : \mathcal{U} \to \mathcal{V}$ is a diffeomorphism. In fact, they can be taken to be geodesically convex (i.e. the geodesic between any x and y in the set is contained in the set), and the curve $t \to \exp_p(tV)$ is also a geodesic.

This exponential can be used to get local coordinates, called godesic normal coordinates, of S near p. Given an orthonormal basis e_1, e_2 of T_pS, there must be an $r \in \mathbb{R}$ such that $x^1e_1 + x^2e_2 \in \mathcal{U}$ for any $(x^1, x^2) \in B_r(0) \subseteq \mathbb{R}^2$, since T_pS with this basis is isomorphic to $\mathbb{R}^2$ through this basis. Then, define $\phi : B_r(0) \to S$ by $\phi(x^1, x^2) = \exp_p(x^1e_1 + x^2e_2)$. This map has the following properties:

- Straight lines through the origin at $B_r(0)$ map to geodesics, and in particular the coordinate axes are geodesics. This isn't necessarily true for lines that don't contain the origin, which can be shown by plugging a line into the geodesic equation.
- The induced metric is Euclidean at the origin in $B_r(0)$: since $[D\exp_p](0) = \mathrm{id}$, then $\frac{\partial\varphi(0,0)}{\partial x^1} = [D\exp_p]_0 \cdot e_1 = e_1$, and similarly for e_2. Since these coordinates are orthonormal, then $g_{ij}(0) = \langle e_1, e_2 \rangle = \delta_{ij}$.
- The Christoffel symbols vanish at the origin in $B_r(0)$: Γ_{ij}^k are given by $\nabla_{E_i}E_j = \sum_k \Gamma_{ij}^k e_k$. By Lemma 11.1, it is possible to express Γ_{ij}^k in terms of the $\frac{\partial g^{ij}}{\partial x^k}$, and this relationship is invertible, so if the former is zero, then the latter must be as well. This has the interesting implication that $\nabla_V V = 0$ for all V at the origin, so let $V = E_i + E_j$. Thus, $\nabla_{E_i+E_j}(E_i + E_j) = 0$, and thus $\nabla_{E_i}E_i + \nabla_{E_i}E_j + \nabla_{E_j}E_i + \nabla_{E_j}E_j = 0$, and therefore $\nabla_{E_i}E_J + \nabla_{E_j}E_i = 0$. Since $\nabla_{E_i}E_j = \nabla_{E_j}E_i + [E_i, E_j]$, but the bracket is zero because they are coordinates. Thus, $\nabla_{E_i}E_j = 0$ at the origin, so the Christoffel symbols vanish.
- Thus, Taylor-expanding the induced metric in the parameter domain has no first-order terms: $g_{ij}(x) = \delta_{ij} + \mathcal{O}(\|x\|^2)$ for $x \in B_r(0)$.

Thus, there are coordinates that are Euclidean to first order at a point. Is it possible to do better, and make the map Euclidean on an entire neighborhood? This is the map-maker's problem, since a solution would imply the existence of a map of a surface without distortion. Formally, the equations necessary to solve this to more than second-order are overdetermined. The following integrability condition has to hold:[22]

$$0 = \frac{\partial\Gamma_{jk}^s}{\partial x^i} - \frac{\partial\Gamma_{ik}^s}{\partial x^j} + \Gamma_{jk}^t\Gamma_{it}^s - \Gamma_{ik}^t\Gamma_{jt}^s$$

for all i, j, k, and s. This encodes the rigidity of the surface, or whether it can be deformed into a map.

[22] A lot of calculations go here. For a full account, consult Spivak.

Definition. The Riemann curvature $(3,1)$-tensor of S is given by $\mathrm{Rm}(X,Y,Z) = \nabla_Y \nabla_X Z - \nabla_X \nabla_Y Z - \nabla_{[X,Y]} Z$.

This can be expanded on a basis as

$$\mathrm{Rm} = \sum_{i,j,k,s} R^s_{ijk} \omega^i \otimes \omega^j \otimes \omega^k \otimes E_s,$$

where

$$R^s_{ijk} = \frac{\partial \Gamma^s_{jk}}{\partial x^i} - \frac{\partial \Gamma^s_{ik}}{\partial x^j} + \Gamma^t_{jk}\Gamma^s_{it} - \Gamma^t_{ik}\Gamma^s_{jt}.$$

In Euclidean space, second partial derivatives commute: $\nabla_{E_i}\nabla_{E_j} = \nabla_{E_j}\nabla_{E_i}$, but for non-Euclidean spaces, the Riemann curvature tensor measures the failure of commutativity of this. This provides some intuition for why it looks the way it does.

Additionally, there is a relation between the very abstract notion of Riemann curvature and the incredibly concrete second fundamental form:

Theorem 15.1 (Theorema Egregium of Gauss[23])**.** *$R^s_{ijk} = A_{jk}A^s_i - A_{ik}A^s_j$, where $A^s_i = \sum_t g^{st}A_{it}$.*

Proof. Let $\overline{\nabla}$ be the covariant derivative in Euclidean space (i.e. ordinary differentiation in $\mathbb{R}^3$) and ∇ is the covariant derivative in S, then $\overline{\nabla}_X Y = (\overline{\nabla}_X Y)^{\parallel} + (\overline{\nabla}_X Y)^{\perp} = \nabla_X Y + A(X,Y)N$ as before. Then, for $i,j,k,\ell = 1$ or 2, take

$$\begin{aligned} 0 &= \left\langle \overline{\nabla}_{E_i}\overline{\nabla}E_j - \overline{\nabla}_{E_i}\overline{\nabla}_{E_j}E_k, E_\ell \right\rangle \\ &= \left\langle \overline{\nabla}_{E_j}\left(\nabla_{E_i}E_k + A(E_i,E_k)N\right) - \overline{\nabla}_{E_i}\left(\nabla_{E_j}E_k + A(E_j,E_k)N\right), E_\ell \right\rangle \end{aligned}$$

Looking specifically at the first term (since the second term is similar):

$$\begin{aligned} \left\langle \overline{\nabla}_{E_j}\left(\nabla_{E_i}E_k + A(E_i,E_k)N\right), E_\ell \right\rangle &= \left\langle \overline{\nabla}_{E_j}(\nabla(E_iE_k) + E_j(A(E_i,E_k))N + A(E_i,E_k)\overline{\nabla}_{E_j}N, E_\ell \right\rangle \\ &= \langle \nabla E_j \nabla_{E_i}E_k, E_\ell\rangle + 0 - A(E_i,E_k)A(E_j,E_\ell). \end{aligned}$$

where the inner product is cancelled out because E_ℓ is a tangent vector, so it is orthogonal to the normal component. Then, the formula follows by plugging in the definitions. ⊠

The Riemann $(4,0)$-tensor can be given by lowering an index: $R_{ijk\ell} = \sum_s g_{\ell s}R^s_{ijk}$. In two dimensions, the Theorema Egregium has only one term $R_{1212} = A_{11}A_{22} - A^2_{12} = \det A$. In an orthonormal basis, this is the product of the principal curvatures, which is also known as the Gauss curvature. Thus, this is an intrinsic quantity!

Definition. Two surfaces S and S' with metrics g and g', respectively are isometric if there is a map $\phi : S \to S'$ such that for any $X_p, Y_p \in T_pS$ and $p \in S$, $g'(D\phi(X_p), D\phi(Y_p)) = g(X_p, Y_p)$.

This just means that the intrinsic geometry is preserved. Rigid motions of surfaces embedded in $\mathbb{R}^3$ are isometries, for example, though other isometries certainly exist: one has the isometry between a catenoid and a helicoid, as in this animation. Thus, one has the notion of extrinsically induces versus intrinsic isometries; another example is a plane versus a non-planar developable surface such as a cylinder: there is a surfact that resembles a cylinder but with deformations so that it isn't extrinsically the same, but has the same intrinsic isometries.

Isometries can also be contrasted with deformation: take a bowling-pin-like surface of revolution and reflect it across the middle, which is an extrinsic isometry. It's a big jump, in some sense, and is considered a distrete isometry. It can also be inverted (which is intrinsic): take the head of the pin and "pop" it inwards.[24]

The take-away message is that isometries are in general rare. The Riemann curvature tensor provides an obstruction, since it must be invariant, but tends not to be.

[23] Latin for "Gauss' Totally Awesome Theorem."

[24] This isn't an isometry for all bowling pins, but one can be constructed for which this works: it must be sufficiently flat near the bottom of the head.